

EXHIBIT B

UNITED STATES DISTRICT COURT
WESTERN DISTRICT OF TEXAS
AUSTIN DIVISION

EVDOKIA NIKOLOVA
Plaintiff,

V.

UNIVERSITY OF TEXAS AT
AUSTIN,
Defendant.

*
*
*
*
*
*
*

CASE NO. 1:19-cv-00877-RP

VIDEOTAPED ORAL DEPOSITION

OF

PETER GLICK, Ph.D.

Friday, November 5, 2021

(Reported Remotely)

VIDEOTAPED ORAL DEPOSITION OF PETER GLICK,
Ph.D., produced as a witness at the instance of the
Defendant, and duly sworn, was taken in the above-styled
and numbered cause on Friday, November 5, 2021, from
9:01 a.m. to 4:02 p.m., before Debbie D. Cunningham,
CSR, in and for the State of Texas, reported remotely
via Machine Shorthand, pursuant to the Federal Rules of
Civil Procedure.

--ooOoo--

<p>22</p> <p>1 case, in particular, so far?</p> <p>2 A. Again, I'd have to go and look; but I think I</p> <p>3 charged a 7500-dollar retainer. I may have, you know,</p> <p>4 expended extra hours and then I would have charged a bit</p> <p>5 more; but I'm guessing. And, again, I'd have to go back</p> <p>6 and look. I didn't prepare financial data here for this</p> <p>7 deposition, but I would think it's under \$10,000.</p> <p>8 Q. You charge \$5,000 -- according to your report,</p> <p>9 you charge \$5,000 per day for deposition and trial</p> <p>10 testimony?</p> <p>11 A. Correct.</p> <p>12 Q. Okay. I'd like to start with Section II of</p> <p>13 your report, called Type of Testimony. In this you</p> <p>14 talk -- in this section you refer to social framework</p> <p>15 testimony; and you use quite a number of terms,</p> <p>16 including stereotype, bias, and discrimination. Can you</p> <p>17 please define what you mean by stereotype?</p> <p>18 A. So by stereotypes -- and just to preface this,</p> <p>19 I'm giving definitions that are kind of the common</p> <p>20 understanding in my field. So stereotypes would be</p> <p>21 traits that people associate with a specific social</p> <p>22 category.</p> <p>23 Q. And am I correct in understanding that those</p> <p>24 traits may be accurate, or they may be inaccurate?</p> <p>25 A. Stereotypes can certainly have, you know,</p>	<p>24</p> <p>1 recall you generally distinguishing between explicit and</p> <p>2 implicit bias in your report. Am I correct in</p> <p>3 understanding that when you generally use the word bias</p> <p>4 in your report, you are not distinguishing between</p> <p>5 explicit and implicit bias?</p> <p>6 A. I cannot recall making those distinctions.</p> <p>7 I'd have to go through my report with a fine-tooth comb,</p> <p>8 but I believe that I did not introduce those</p> <p>9 distinctions in the report.</p> <p>10 Q. You also introduce other types of biases or</p> <p>11 other ways to categorize biases, and the explicit versus</p> <p>12 implicit was one of those. What are other types of ways</p> <p>13 to categorize biases?</p> <p>14 A. Actually, I didn't really have anything in</p> <p>15 mind when I said that. Sorry. You know, implicit</p> <p>16 versus explicit is a -- you know, is a -- that's what's</p> <p>17 generally talked about and researched in my field. I</p> <p>18 don't have anything off the top of my head that I can</p> <p>19 think of of other kinds of categories of bias or some</p> <p>20 other kind of distinction in mind.</p> <p>21 Q. What do you mean when you use the word</p> <p>22 discrimination in your report?</p> <p>23 A. Discrimination refers to treating people</p> <p>24 differently based on their social category memberships.</p> <p>25 Q. When you use that social category or</p>
<p>23</p> <p>1 differing degrees of accuracy.</p> <p>2 Q. What do you mean by bias?</p> <p>3 A. So bias would be leanings toward more or less</p> <p>4 favorable kinds of impressions or actions toward</p> <p>5 individuals based on their social category membership.</p> <p>6 Q. And when you use the word bias, does that</p> <p>7 include unconscious bias?</p> <p>8 A. Well, there's certainly different kinds of</p> <p>9 categories that we could talk about of bias. In my</p> <p>10 field we do talk about explicit versus implicit bias</p> <p>11 would be more the terms that might correspond with what</p> <p>12 you're talking about.</p> <p>13 Q. So when you -- so when you use the term bias</p> <p>14 in your report, are you including both explicit and</p> <p>15 implicit bias when you use that term?</p> <p>16 A. Bias would be an overall umbrella term.</p> <p>17 Q. So is the answer "yes"?</p> <p>18 A. That would be yes, it could be one or the</p> <p>19 other.</p> <p>20 Q. Or it could be both?</p> <p>21 A. Yeah, sorry. I was about to finish. Or it</p> <p>22 could be both.</p> <p>23 Q. Sorry. I didn't mean to interrupt you.</p> <p>24 A. No, no, that's fine.</p> <p>25 Q. In your report do you distinguish -- I don't</p>	<p>25</p> <p>1 membership term, can you please explain what that means?</p> <p>2 A. So we categorize -- people categorize others</p> <p>3 on a whole variety of dimensions. Things like gender is</p> <p>4 a category that people very frequently, commonly put</p> <p>5 people in, put other people into. You could talk</p> <p>6 about things like ethnicity or race. Social class could</p> <p>7 be used as a category. So, really, those -- anything</p> <p>8 that -- that distinguish people and put them into groups</p> <p>9 that are perceived to be somewhat discreet. That is</p> <p>10 different kinds of groups can be used as a basis of</p> <p>11 categorization.</p> <p>12 Q. And I believe you, also, in your report use</p> <p>13 the word prejudice. What does prejudice mean to you?</p> <p>14 A. Well, more commonly in my field prejudice is</p> <p>15 generally associated with the differential emotions</p> <p>16 towards different groups. That's how it would be</p> <p>17 commonly used. Again, it's kind of an umbrella term.</p> <p>18 Q. Okay. So does prejudice incorporate a concept</p> <p>19 of intentional or conscious feelings about a particular</p> <p>20 group?</p> <p>21 A. So, you know, first off, that term</p> <p>22 intentional, that, I think, has a lot of legal meanings</p> <p>23 and may have different interpretations. I think intent</p> <p>24 is a concept that's used very commonly in legal</p> <p>25 language; although, again, I want to put the caveat out</p>

26

1 there that I am not a lawyer. So, normally, intent
2 isn't exactly the way that people in my field tend to
3 put this. And when we talk about implicit bias, we're
4 talking more about whether people are aware of their
5 motivations and aware of the mental processes that
6 might, for example, lead them to treat people
7 differently based on their category memberships.

8 Q. So how does that -- how does that implicit
9 bias or un- -- being unaware affect discrimination?

10 (Simultaneous speakers.)

11 A. All right. So I --

12 Q. Go ahead. I'm sorry.

13 A. Oh, I'm sorry. Sorry. Go ahead. No, finish,
14 please.

15 Q. No, go ahead.

16 A. All right. So you're asking me sort of --

17 kind of an example of implicit bias? Is that what
18 you're asking me?

19 Q. No. I'm trying -- my apologies. I'm trying
20 to understand the link between implicit bias and
21 discrimination. So let me ask you this question: Can
22 implicit bias lead to discrimination as you define that
23 term?

24 A. Yes.

25 Q. And implicit bias, meaning a bias that someone

28

1 A. Yeah, we should -- that was a convoluted
2 question to me.

3 Q. I'll ask it a different way.

4 A. Yeah, yeah, yeah.

5 Q. Does a person have to be aware of their bias
6 to engage in discrimination on the basis of that
7 particular social category?

8 A. Theoretically, there's the possibility that
9 they don't have to be. Can we prove that they're
10 completely unaware? That's -- that's not -- you know,
11 the research doesn't necessarily show that people are
12 completely unaware of their biases.

13 Q. Can you -- do you prove that they're aware
14 when you use the word discrimination? Do you -- is that
15 first -- is proving that they're aware of their bias a
16 condition upon using the word discrimination?

17 A. Overall, I would say no.

18 Q. Okay. So -- and is it true that people have
19 biases that they are not aware of?

20 A. Theoretically, that's possible. What I'm
21 saying is proving that that's true is really quite
22 tricky or impossible because people, when you're trying
23 to assess whether they're aware, that relies on them
24 reporting their awareness; and there are well-known
25 biases in reporting that awareness.

27

1 is not even aware they have, can lead to differential
2 treatment as, i.e., discrimination based on a membership
3 in a social category as you define it?

4 A. So the way that we distinguish between
5 explicit and implicit bias concerns methodologies where
6 I don't think we can't pin down that the person is
7 completely unaware of their bias because, with explicit
8 bias, you're asking them the question outright; and
9 people have a lot of motives not necessarily to tell
10 you -- you know, to reveal that they -- that they have a
11 bias. So there are social desirability concerns when
12 you ask somebody directly about biases toward different
13 groups. The tendency is to deny that.

14 So implicit bias research kind of
15 developed as a way to bypass those social desirability
16 tendencies by creating measures where people -- people's
17 responses aren't as deliberately controlled, that is,
18 they can't as easily suppress biases that they may have.
19 So the question of whether they're aware of that
20 implicit bias is a very tricky question that I think we
21 can't completely pin down.

22 Q. So when you use the word discrimination, you
23 are not using that word to mean discrimination based on
24 a bias of which someone is necessarily even aware of; is
25 that correct?

29

1 Q. But you can -- so put in a different way, you
2 can -- social science has been able to identify biases
3 that affect behavior without being able to prove that
4 the individual is aware that they have the bias?

5 A. Well, so there is research, for example, where
6 people use what is called a bogus pipeline, basically,
7 convince people that you have a lie detector sort of
8 arrangement, where if they don't tell the truth, they're
9 going to be exposed. And what you typically find in
10 those circumstances is that people will admit to biases
11 more readily; but, you know, in research that doesn't do
12 that -- and that's not -- you know, that's a
13 labor-intensive method that not everybody employs. In
14 that research that does not use that kind of
15 methodology, we can't tell for sure the degree of
16 awareness or lack of awareness that a person might have.

17 So, theoretically, yes. Theoretically, a
18 person might not be aware of their tendency to treat
19 people differently or they may not acknowledge it to
20 themselves, which is maybe slightly different. They may
21 be not acknowledging it deliberately to other people
22 because they don't want to be -- you know, they don't
23 want to see themselves as biased; and they don't want
24 other people to see themselves as biased. So there can
25 be kind of a suppression of that when you ask people

<p>30</p> <p>1 about it.</p> <p>2 So, you know, what I'm saying is: We</p> <p>3 can't prove lack of awareness. Okay? We can show that</p> <p>4 people who deny that they're being discriminatory and</p> <p>5 who claim to have very strong motives to be non-</p> <p>6 discriminatory, we can demonstrate that those people</p> <p>7 may, nevertheless, discriminate; but we don't know their</p> <p>8 degree of awareness.</p> <p>9 Q. And so in your report when you use the words</p> <p>10 bias and discrimination, I presume that you are not</p> <p>11 using those terms to presume that the person is</p> <p>12 necessarily even aware that they have those biases or</p> <p>13 are engaging in what you call discrimination?</p> <p>14 A. I guess the way I would put it is in my report</p> <p>15 I'm not making claims of people's degree of awareness of</p> <p>16 their biases.</p> <p>17 Q. And just to be clear, when you use the word</p> <p>18 bias, you're not making any claim that that person is</p> <p>19 aware of the bias?</p> <p>20 MR. SCHMIDT: Objection, form.</p> <p>21 A. I'm not making any claim about people's</p> <p>22 awareness of their biases.</p> <p>23 Q. (BY MR. GIBSON) And you are not necessarily</p> <p>24 making a claim that the decision based on that bias is</p> <p>25 driven by an intent or -- to discriminate against that</p>	<p>32</p> <p>1 describing as discriminating has an intent to</p> <p>2 discriminate?</p> <p>3 A. I'm not requiring that the person will report</p> <p>4 that they have an intent to discriminate. I'm not</p> <p>5 requiring that they are necessarily thinking that they</p> <p>6 are discriminating.</p> <p>7 Q. How would you describe animus? Are you</p> <p>8 familiar with the term animus?</p> <p>9 A. Animus. Is that -- are you asking me from a</p> <p>10 legal perspective, or are you asking me from a</p> <p>11 psychological perspective?</p> <p>12 Q. I'm asking you from a -- your understanding as</p> <p>13 an expert in this case. I'm not asking you for a legal</p> <p>14 opinion. I'm asking you: What do you -- what do you</p> <p>15 understand that word to mean --</p> <p>16 A. Well, it's not --</p> <p>17 Q. -- in the context of bias and discrimination?</p> <p>18 A. Right. Okay. Well, that specific term is not</p> <p>19 commonly used in my area; but my understanding of -- I</p> <p>20 can give you my dictionary definition of the term animus</p> <p>21 as far as I understand it, which would be hostility or</p> <p>22 antipathy toward -- in the context of discrimination,</p> <p>23 hostility or antipathy toward a group and members of</p> <p>24 that group by virtue of their membership.</p> <p>25 Q. And, again, just to make sure I'm clear, your</p>
<p>31</p> <p>1 particular group?</p> <p>2 A. Well, again, I mean, people reporting on</p> <p>3 intent is very tricky in research because people,</p> <p>4 especially in the context of a legal case, are unlikely</p> <p>5 to say: Yes, I intentionally discriminated, so. But</p> <p>6 even in an -- you know, in an experiment or research</p> <p>7 where you ask them, people are commonly suppressing</p> <p>8 intent to discriminate. So, you know, I'm not making</p> <p>9 claims about people's awareness of their intent to</p> <p>10 discriminate.</p> <p>11 Q. Are you making claims about their intent to</p> <p>12 discriminate?</p> <p>13 A. I think I said no to that question.</p> <p>14 Q. I just want to make sure.</p> <p>15 A. Yeah.</p> <p>16 Q. So you're not -- you're not -- when you use</p> <p>17 the word discrimination, you are not making -- you are</p> <p>18 not including in that definition any -- you're not</p> <p>19 presuming or requiring that there be an intent to</p> <p>20 discriminate when you use the word discriminate?</p> <p>21 MR. SCHMIDT: Objection, form.</p> <p>22 A. Could you just repeat that?</p> <p>23 Q. (BY MR. GIBSON) When you use the word</p> <p>24 discriminate or discrimination in your report, you are</p> <p>25 not including a requirement that the person you're</p>	<p>33</p> <p>1 use of the word discrimination in your report does not</p> <p>2 require that the actions be motivated by hostility or</p> <p>3 antipathy to that group; is that correct?</p> <p>4 MR. SCHMIDT: Objection, form.</p> <p>5 A. So, more broadly, discrimination can take</p> <p>6 place in the absence of -- of, you know, overt animus or</p> <p>7 antipathy. So that is -- that is a kind of</p> <p>8 discrimination that can occur.</p> <p>9 Q. (BY MR. GIBSON) And am I correct that your</p> <p>10 report does not distinguish -- when it uses the word</p> <p>11 discrimination, it does not distinguish between</p> <p>12 discrimination that is premised upon overt hostility or</p> <p>13 antipathy versus discrimination that may be based on an</p> <p>14 implicit bias?</p> <p>15 A. Yeah. So I think the answer to that is "yes."</p> <p>16 I'm not making my -- say, discrimination more generally.</p> <p>17 It could include both of those things. You know, so</p> <p>18 there's a variety of things that could be included under</p> <p>19 that umbrella term that lead to unfair discriminatory</p> <p>20 treatment toward people because of a group membership.</p> <p>21 Q. Okay. You, also, in your report at various</p> <p>22 points describe how bias and discrimination are very</p> <p>23 context specific. What do you mean by that?</p> <p>24 A. So what I mean by that is that -- that,</p> <p>25 really, there are circumstances that are more likely to</p>

34

1 promote the possibility of bias that we know from the
2 research that under certain circumstances people will
3 engage in bias more readily. So imagine a person who
4 does have a bias, has a stereotype that leads to bias or
5 discrimination or can lead to bias or discrimination.
6 That person, under certain circumstances, may not, you
7 know, allow that bias to affect their, say, evaluation
8 of another person.

9 So some circumstances permit bias to
10 occur more readily. Some circumstances, you know,
11 actively kind of encourage the expression of bias.
12 Other circumstances inhibit or mitigate the likelihood
13 of bias, and we know this from the research.
14 So it's -- I mean, you can think of an
15 extreme case where somebody is so -- you know, so
16 enamored of a hateful ideology that they're always going
17 to express their bias no matter what cost, no matter
18 what situation. That's not the typical kind of ways in
19 which bias gets expressed. That's not a typical kind of
20 case.

21 Q. Does cultural context matter with respect to
22 whether or not biases may lead to discriminatory
23 decisions?

24 MR. SCHMIDT: Objection, form.

25 A. That's an extremely broad statement.

35

1 Q. (BY MR. GIBSON) Question. Question, not a
2 statement.

3 A. Oh, okay. An extremely broad, broad question.
4 You know, at a very broad level, cultural context
5 certainly can matter; but I'm not sure exactly what you
6 have in mind or exactly what you mean by cultural
7 context. So I'm a little wary of this question unless I
8 understand a little more in detail what you're -- what
9 you're getting at as an example.

10 Q. Sure. Well, let's start by talking about
11 location. Does, for instance, cultural context in
12 Saudi Arabia matter relative to the cultural context in
13 the United States as to whether bias against, oh, let's
14 say on the basis of sex may be expressed by a particular
15 individual?

16 A. Absolutely, perceived social norms will, you
17 know, affect the likelihood of expression of prejudice;
18 but what I want to add to this is that what happens when
19 people perceive the cultural norms to be prohibiting the
20 expression of prejudice is that people will seize on any
21 other reasons or excuses that seem non-prejudicial to
22 express their prejudices.

23 So with all of these things, when I say
24 some factor matters, it's not the only factor; and it
25 also can interact with other things going on in a

36

1 situation. So to say that, you know, I wouldn't make an
2 absolute statement, for instance, that, oh, in a culture
3 that suggests prejudice and bias is a bad thing, that
4 nobody is going to act on a bias and discriminate.
5 That's not true. All right? It means that
6 discrimination becomes more subtle. It becomes more
7 covert. It becomes more expressed in situations where
8 people have a sense that they can deny that they have
9 discriminatory motives. So that's really what I mean by
10 it's highly contextual. It's complicated.

11 Q. It sounds like you were referring to a
12 particular individual, meaning it's an individual who
13 has a bias in one -- if they're placed in one
14 environment may express that bias more overtly versus if
15 they're placed in a different environment, they may
16 express that bias more covertly?

17 A. That's a good example, yes. That's what I was
18 getting at as an example.

19 Q. Is it also true the context itself can affect
20 the research such that whether or -- like, is it
21 important that the research to determine whether or not
22 a particular bias is affecting a particular situation,
23 such as a workplace, that the research is context
24 specific to the relevant context, meaning, does research
25 about workplaces in oil rigs in Saudi Arabia, about the

37

1 treatment of women on oil rigs in Saudi Arabia, does it
2 matter that that's on oil rigs in Saudi Arabia when
3 we're talking about whether that's relevant to white-
4 collar workers in New York City; or is it important that
5 the research be more closely aligned with white-collar
6 workers in New York City if you're going to make the
7 statement that the research applies to the specific
8 context in which you're seeking to apply it?

9 A. Okay. I think I -- I get where you're going
10 with this and what you're trying to get at. And just --
11 so this, you know, requires that you really have a kind
12 of deeper level of understanding of how the field of
13 social psychology works. And so what we try to do
14 generally is to try to understand variables that -- that
15 influence these kinds of things. What are the kinds of
16 factors that influence people's tendency to discriminate
17 and understand that -- understand those as psychological
18 variables.

19 So I'll just -- I'm going to try to give
20 an example. So let's say that you're trying to
21 understand how forest fires work, right? You could
22 still study that process outside the context of a
23 forest. Let's say you're studying the role of oxygen in
24 fueling a fire, right? Oxygen is, you know, a factor
25 that plays into whether the fire is going to be

<p>46</p> <p>1 makers about relevant, scientifically validated</p> <p>2 principles. This is not the same as performing</p> <p>3 diagnostic tests for individuals or organizations named</p> <p>4 in a legal case."</p> <p>5 A. Correct.</p> <p>6 Q. To be clear, your report does not purport to</p> <p>7 provide scientific analysis of the facts of this case;</p> <p>8 is that correct?</p> <p>9 A. Well, I wouldn't put it that way, no. I would</p> <p>10 say that the Application to the Case section is informed</p> <p>11 by expertise in the scientific framework. So that</p> <p>12 expertise in the scientific framework is informative in</p> <p>13 identifying areas where behavior kind of meshes with or</p> <p>14 potentially matches the kind of pattern you would expect</p> <p>15 if certain kinds of bias were occurring. It also</p> <p>16 identifies contextual factors that, you know, might</p> <p>17 mitigate or make bias more likely.</p> <p>18 So it's based on the scientific</p> <p>19 framework, but I want to make clear the distinction</p> <p>20 that, in the application to the case, you're not</p> <p>21 conducting a study specifically on the organization or</p> <p>22 the people involved, which would, I think, be an undue</p> <p>23 burden and also problematic from a scientific</p> <p>24 perspective; but instead, you're using the scientific</p> <p>25 framework and expertise to identify these kinds of</p>	<p>48</p> <p>1 A. I am very clear in my report that I am not</p> <p>2 conducting studies on the organization, University of</p> <p>3 Texas, or the specific individuals involved and,</p> <p>4 therefore, cannot come to a scientific conclusion about</p> <p>5 whether or not discrimination occurred in this case.</p> <p>6 And I'm just very careful about that and very honest</p> <p>7 about that.</p> <p>8 Q. You do provide opinions that the facts of this</p> <p>9 case are consistent with discrimination, don't you?</p> <p>10 A. I point out where the -- there are, for</p> <p>11 instance, contextual factors that will make</p> <p>12 discrimination more or less likely; and I can talk</p> <p>13 about specific examples in this case if you like. And</p> <p>14 I -- and I point out where, based on the research, we'd</p> <p>15 expect a certain pattern of kinds of judgments and</p> <p>16 some -- and that, you know, seem to match with -- match</p> <p>17 with the -- some of the facts of the case or some of the</p> <p>18 testimony in the case.</p> <p>19 So it's really kind of a -- I'm pointing</p> <p>20 out what patterns we would expect based on the research,</p> <p>21 when would discrimination be more or less likely to</p> <p>22 occur, which form would it be more or less likely to</p> <p>23 take, and then point out areas for the jury to consider</p> <p>24 without telling them -- I'm, again, very explicit that</p> <p>25 I'm not ultimately making the conclusion. They need to</p>
<p>47</p> <p>1 factors that are, you know, potentially involved in the</p> <p>2 case.</p> <p>3 Q. You do not apply a scientifically valid</p> <p>4 methodology to provide the opinions that you provide in</p> <p>5 the section called Application to the Case; is that</p> <p>6 correct?</p> <p>7 A. Well, again, my opinions are informed by --</p> <p>8 Q. That wasn't my question. I'm not asking -- we</p> <p>9 can all agree that your opinions -- my guess -- well,</p> <p>10 let me ask you this question: Do you have any opinions</p> <p>11 about discrimination in the workplace that are not</p> <p>12 informed by your background? Can you -- can you cast</p> <p>13 off your scientific background when you -- cast that off</p> <p>14 when you provide opinions about discrimination in the</p> <p>15 workplace?</p> <p>16 A. I don't think so, no.</p> <p>17 Q. Right. So every opinion that you provide</p> <p>18 about discrimination in the workplace is informed by the</p> <p>19 fact that you're a scien- -- a social scientist?</p> <p>20 A. It's informed by my understanding of research</p> <p>21 with a known error rate and all that, yes.</p> <p>22 Q. Right. So we -- we all understand that. My</p> <p>23 question is: Did you conduct any scientific study of</p> <p>24 the facts of this case to come to the opinions that you</p> <p>25 provide in the section called Application to this Case?</p>	<p>49</p> <p>1 make a legal judgment. I can't make a scientific</p> <p>2 judgment that says, you know, that, with a certain</p> <p>3 degree of probability, that discrimination did or did</p> <p>4 not occur.</p> <p>5 Q. For instance, you say, "The Dean's decision-</p> <p>6 making about Dr. Nikolova is consistent with bias</p> <p>7 towards pregnant women, mothers, and workplace</p> <p>8 accommodation policy use." That's on page 57 if you</p> <p>9 would like to refer to it.</p> <p>10 A. Yeah, I recall that. That's correct.</p> <p>11 Q. So you are making a statement that the Dean's</p> <p>12 decision-making is consistent with bias towards pregnant</p> <p>13 women?</p> <p>14 A. Yeah, I'm making a statement that if there</p> <p>15 were bias towards --</p> <p>16 Q. I'm sorry. That's not the words -- please</p> <p>17 don't restate what you said.</p> <p>18 A. Okay.</p> <p>19 Q. You didn't --</p> <p>20 (Simultaneous speakers.)</p> <p>21 Q. You said, "The Dean's decision-making about</p> <p>22 Dr. Nikolova is consistent."</p> <p>23 A. Yeah. Yes, yes. I'm not denying anything.</p> <p>24 Q. Great. Great.</p> <p>25 A. I stand by that statement. I'm just</p>

<p>50</p> <p>1 explaining to you, you know, how I conceive of that 2 statement; and it's in the context of saying that this 3 is not a scientifically -- this is not a conclusion that 4 has a specific probability value to it and that the 5 ultimate decision about whether it was, indeed, 6 discrimination is up to the case decision-makers. 7 Q. Is there any scientific methodology that you 8 used that's supported by a peer-accepted and peer- 9 reviewed methodology that allows you to say that 10 Dr. Nikolova's de- -- I'm sorry -- the Dean's decision- 11 making about Dr. Nikolova is consistent with bias toward 12 pregnant women, mothers, and workplace accommodation 13 policy use? 14 MR. SCHMIDT: Objection, form. 15 A. Yeah, could you repeat or restate that? 16 Q. (BY MR. GIBSON) Sure. Is there a scientific 17 theory or technique that is generally accepted in the 18 scientific community that supports the statement that 19 says, "The Dean's decision-making about Dr. Nikolova is 20 consistent with bias toward pregnant women, mothers, and 21 workplace accommodation policy use"? 22 A. Again, if you're asking if I conducted a 23 scientific study on the organization that could draw 24 that conclusion to a scientific certainty, the answer is 25 "no."</p>	<p>52</p> <p>1 Q. And you mentioned that you have served as an 2 expert in the Gould v. Interface case. 3 A. Correct. Yes, I did. 4 Q. In that report did you include a section in 5 that report called Application to the Case or a similar 6 type of section where you sought to apply social science 7 to the facts of that case? 8 A. Yes, I did. 9 Q. In your report you also talk about acceptance 10 of social framework testimony by Courts, again, in 11 Section II. Are you seeking to provide an opinion 12 about -- a professional opinion, an expert opinion, on 13 the acceptance of social science framework testimony in 14 legal cases? 15 A. Again, I'll restate that I am not a lawyer. 16 Okay? So -- 17 Q. That wasn't my question. 18 A. All right. That's my preface to my answer. 19 That wasn't my complete answer. All right? You said 20 you wouldn't interrupt me, so I wasn't done. 21 All right. So -- oh, now, can you 22 restate the question because now I'm -- I'm losing the 23 question? 24 Q. Are you providing -- your report makes 25 statements about the acceptance of social framework</p>
<p>51</p> <p>1 Q. Are there any peer-reviewed articles that -- 2 where researchers engage in the type of statements about 3 an individual case based on a re- -- the limited review 4 of evidence that you conducted in this case to make 5 statements about whether or not conduct is consistent 6 with bias? 7 A. Okay. So, here, I feel like what you're doing 8 is you're saying, you know, would I submit the last 9 section of this report to a journal as though it's a 10 journal article that would be peer reviewed and would be 11 published. No, that's not the intent of that section. 12 That's not the context of that section. My 13 understanding is that this is a legal case; and that is 14 the rendering of an expert opinion, again, rooted in the 15 science of discrimination from somebody who's an expert 16 in it. 17 I believe that other colleagues who were 18 similarly rooted and understand this research would come 19 to similar conclusions, right? But is this the same as 20 doing a scientific study? No. I'm very clear about 21 that. I'm not trying to deny anything here. I feel 22 like you're trying to frame that question in a way that, 23 you know -- that makes it sound worse or something like 24 that. But, again, I was very straightforward about my 25 framing of this section and the context of this section.</p>	<p>53</p> <p>1 testimony in legal cases by Courts. Do you have -- are 2 you qualified to make that statement? 3 A. The statement that I made about that is based 4 on my experience. I've done, I don't know, maybe two 5 dozen cases. My testimony, others have challenged -- 6 you know, tried to get my testimony excluded. The 7 opposing side has tried to get my testimony excluded; 8 and with the exception of the case that you're well 9 aware of that was a recent case, my testimony has not 10 been excluded. And, in fact, in a federal court in 11 Boston the judge excluded other experts but denied the 12 motion to exclude me and specifically praised the way in 13 which I structured my report and drew my conclusions, 14 which is similar to the way I did it in this case. 15 And, you know, social framework testimony 16 has become a bone of contention, you know; and judges, 17 in my understanding -- I'm, again, not claiming to be an 18 expert in the legal treatment of this testimony. My 19 understanding is that the rules have generally permitted 20 what I do, have even -- and in one case even praised 21 exactly how I do it; but the rules tend to shift from 22 judge to judge. They seem to have a lot of discretion 23 to interpret those rules. 24 And so, yeah, I would love it if the 25 rules were so clear cut and the application of the rules</p>

<p>58</p> <p>1 never been excluded. This sort of testimony has been 2 accepted. So in my experience, it's whatever, 1 out of 3 25; and it was a recent decision. 4 I am not claiming to have the expertise 5 to make this as a legal argument. As far as the 6 scholars, Monahan and Walker, they don't get to make the 7 rules. The Courts make the rules. The Courts, in my 8 experience, have generally accepted this kind of 9 testimony; and if a judge doesn't, there's not anything 10 I can really do about it. 11 Q. I would like to move to Section III of your 12 report. 13 A. Sure. 14 Q. Page 9. 15 A. Page -- okay. Yeah. 16 Q. I'll just -- is this a list -- this says this 17 is a list of the documents you reviewed in connection 18 with drafting this report; is that correct? 19 A. That is correct. 20 Q. How did -- how were these documents 21 identified, to your knowledge? 22 A. These documents were sent to me by 23 Mr. Schmidt. 24 Q. So these were documents provided to you by 25 Plaintiff's Counsel?</p>	<p>60</p> <p>1 regarding the facts of the case? 2 A. The additional information that I have 3 received, I did receive your Motion for Summary 4 Judgment, I believe, and then the response to that 5 motion. 6 Q. Okay. Did you -- did you look at the exhibits 7 to that? 8 A. I did not look at the exhibits to that. I did 9 not receive those exhibits. 10 Q. Okay. So you saw the motion but not the 11 actual evidence cited that was cited as part of the 12 motion? 13 A. I did not see the exhibits for those motions. 14 Q. Okay. Just to confirm, you have not reviewed 15 the deposition of Dean Wood in this case; is that 16 correct? 17 A. That's correct, I haven't reviewed her 18 deposition; but I, you know, have reviewed her 19 explanation of her decision-making or the Committee's -- 20 Committee on Academic Freedom's notes on that. 21 Q. Are those -- so you reviewed notes taken by 22 other people about what she purportedly said to them? 23 A. Yes, I reviewed notes of what a Committee 24 contemporaneously documented about what Dean Wood said 25 to them. I also -- the Dean's assessment letter for</p>
<p>59</p> <p>1 A. Yes, that's usually how it works. 2 Q. Did you ask him for any additional documents 3 after he provided you these documents? 4 A. I asked him to provide me with all of the 5 relevant documents; and usually -- I can't recall 6 specifically -- but, usually, I make it clear that I 7 don't want to be surprised in a deposition by the -- you 8 know, the existence of other documents I should have 9 been aware of. 10 Q. This list does not include any deposition 11 testimony; is that correct? 12 A. I believe the depositions or most of those 13 depositions were not taken at the time that my report 14 was due. So there was, I guess, an agreed-upon 15 deadline, as far as I understand it, as to when my 16 report was due; and depositions had not yet been taken. 17 Q. Have you asked for them since that time? 18 A. Well, I -- I didn't -- I did discuss with, you 19 know, Mr. Schmidt whether I should be reviewing 20 depositions. The report is based on documents that were 21 available at the time of my report, and I very 22 specifically say in the report that I reserve the right 23 to alter opinions based on further information. 24 Q. And my question is: Have you received any 25 additional information since writing your report</p>	<p>61</p> <p>1 Dr. Nikolova. So there were some direct -- you know, 2 direct primary documents, I would call them, that -- 3 that represent Dean Wood's justifications and opinion. 4 Q. Did you review any -- and, again, you didn't 5 review any other deposition testimony in this case, 6 correct? 7 A. Looking at the list, I can't remember any -- 8 any specific depositions in review; but the list -- the 9 list should be complete about what I reviewed at the 10 time that the report was due that was available. And so 11 the list on page 9 of my report should be complete; and 12 then, since then, as I said, I saw the summary motion -- 13 the Motion for Dismissal and the response. 14 Q. Did you review any policies and procedures 15 from the University of Texas? I don't see any listed 16 here. 17 A. Other than what's listed on the list, no. 18 Q. Okay. Did you review any data regarding 19 teacher evaluations and tenure and promotion from the 20 University of Texas regarding the treatment of faculty 21 members, including male versus female faculty members, 22 in those areas? 23 A. Again, I reviewed what was on the list. I 24 think there were definitely some things on the list that 25 discussed teaching evaluations; but -- but other than</p>

<p style="text-align: right;">62</p> <p>1 that, no.</p> <p>2 Q. And as I recall in your report, you presume</p> <p>3 certain facts stated in the Plaintiff's complaint are</p> <p>4 true; or you spec- -- you know, you say, "Assuming these</p> <p>5 facts are true," and then you make statements about</p> <p>6 those facts. Did you --</p> <p>7 A. I -- in my report, I'm careful, trying -- I</p> <p>8 believe I tried not to make credibility judgments; and</p> <p>9 that's another understanding that I have of the rules</p> <p>10 for social framework testimony based on what the Courts</p> <p>11 allow. And my understanding is that I'm not supposed to</p> <p>12 make credibility judgments, that is, who's -- if there's</p> <p>13 a he said/she said, I'm not judging who -- who was more</p> <p>14 accurate, right? So in those cases, as I very carefully</p> <p>15 state in my report, I say if the jury -- something like,</p> <p>16 along the lines of: If the jury finds this credible,</p> <p>17 then, you know, here's something to consider, right? So</p> <p>18 I make that contingency very clear. I'm trying to be</p> <p>19 careful and fair in that kind of statement.</p> <p>20 Q. Do you conduct any scientific research where</p> <p>21 you rely solely on data provided by the subject to</p> <p>22 make -- to provide scientific determinations regarding</p> <p>23 the treatment of that subject, meaning, would you ever</p> <p>24 rely on the person who's alleging to be discriminated</p> <p>25 against in science to give you the evidence -- solely</p>	<p style="text-align: right;">64</p> <p>1 your discussion section to note the potential weaknesses</p> <p>2 of this and caveats about what you can conclude from it.</p> <p>3 Q. Have you ever personally conducted science</p> <p>4 where you make a determination of whether or not someone</p> <p>5 is more or less likely to be subject to bias or</p> <p>6 discrimination based solely on evidence provided by that</p> <p>7 person?</p> <p>8 A. Not that I can think --</p> <p>9 Q. Okay.</p> <p>10 A. -- of offhand, no.</p> <p>11 Q. Can relying on evidence solely by a person who</p> <p>12 claims to have been discriminated against of the alleged</p> <p>13 discrimination introduce bias into a scientific study?</p> <p>14 MR. SCHMIDT: Objection, form.</p> <p>15 A. Again, I'm -- I'm having a hard time decoding</p> <p>16 your question.</p> <p>17 Q. (BY MR. GIBSON) Do you think it's -- do you</p> <p>18 think that it's reflective of bias that you only</p> <p>19 reviewed evidence that was specifically identified and</p> <p>20 told to you that you should review by Plaintiff's</p> <p>21 Counsel, who is an advocate for Plaintiff in this case?</p> <p>22 MR. SCHMIDT: Objection, form.</p> <p>23 A. Again, I asked for whatever would be relevant</p> <p>24 to -- to making this case; and that's a pretty common --</p> <p>25 that's commonly how things proceed. And the -- I will</p>
<p style="text-align: right;">63</p> <p>1 that person to give you evidence of whether or not they</p> <p>2 have been discriminated against?</p> <p>3 MR. SCHMIDT: Objection, form.</p> <p>4 A. Okay. So I think you're kind of -- I'm</p> <p>5 worried that this question is kind of blending science</p> <p>6 and legal determinations, right?</p> <p>7 Q. (BY MR. GIBSON) I don't want to blend the</p> <p>8 two.</p> <p>9 A. All right.</p> <p>10 Q. I want to focus on science. Let's focus on</p> <p>11 science.</p> <p>12 A. Right. So, you know, there certainly might be</p> <p>13 studies where -- I'm sure there are studies where people</p> <p>14 look at the perceptions of individuals on whether</p> <p>15 they're discriminated against; and then in -- you know,</p> <p>16 when you would -- in your discussion section you would</p> <p>17 make clear what the potential drawbacks to relying on</p> <p>18 that as, you know, whether or not discrimination</p> <p>19 occurred.</p> <p>20 So you might be studying people's</p> <p>21 perceptions. That's a legitimate area of psychological</p> <p>22 study. And you might try to correlate those perceptions</p> <p>23 with other indicators of whether there was</p> <p>24 discrimination or not. But, sure, that might happen in</p> <p>25 a scientific study; but you would be very careful in</p>	<p style="text-align: right;">65</p> <p>1 just point out that this -- you know, this evidence, the</p> <p>2 things that I reviewed, include the -- you know, the</p> <p>3 Budget Council's procedures and findings; the UT</p> <p>4 committee who reviewed the decision-making, their</p> <p>5 findings; includes Dean Wood's assessment. So it's not</p> <p>6 like there's only, you know, Dr. Nikolova's voice here.</p> <p>7 This is -- you know, there are primary documents that</p> <p>8 document the ways in which decisions were made.</p> <p>9 And, furthermore, I do not make the</p> <p>10 ultimate conclusion about whether discrimination</p> <p>11 occurred or not. So I feel like your question is kind</p> <p>12 of presuming some things that I didn't do.</p> <p>13 Q. Well, you make opinions about whether or not</p> <p>14 the actions of a dean is consistent with discrimination,</p> <p>15 correct?</p> <p>16 A. I make -- I render opinions about whether</p> <p>17 deans -- you know, the kinds of judgments Dean Wood</p> <p>18 make -- made are more or less likely to -- to have -- to</p> <p>19 be -- you know, to -- that discrimination would be more</p> <p>20 or less likely on those kinds of judgments and also</p> <p>21 whether the pattern of judgments made are consistent</p> <p>22 with the research on, for instance, bias against mothers</p> <p>23 or people who use leave policies for caretaking. So</p> <p>24 those are the kind of judgments I make. Again, I don't</p> <p>25 ultimately render the conclusion about whether</p>

<p>66</p> <p>1 discrimination occurred.</p> <p>2 Q. That wasn't my question.</p> <p>3 So you -- are you comfortable making --</p> <p>4 rendering opinions about whether or not someone's</p> <p>5 conduct is consistent with discrimination based solely</p> <p>6 on being -- looking at evidence provided to you by the</p> <p>7 Plaintiff's attorney?</p> <p>8 A. Well, again, if the -- if the information</p> <p>9 covers the relevant decision-making processes and is --</p> <p>10 includes primary documents, I'm comfortable with those.</p> <p>11 I also very clearly state in my report that if there's</p> <p>12 additional information, that -- that -- you know, that I</p> <p>13 reserve the right to alter my judgment after carefully</p> <p>14 considering it. So if there's some specific sorts of</p> <p>15 information that you would like me to review, I'd be</p> <p>16 happy to review that; but I'm not going to make an</p> <p>17 off-the-cuff, you know, determination. I'd have to</p> <p>18 weigh it with the rest of the information that I have.</p> <p>19 Q. Did you speak with Dr. Nikolova about this</p> <p>20 case directly? Have you ever spoken to her directly?</p> <p>21 A. No, I believe I have not.</p> <p>22 Q. Okay. What is -- in some of your other</p> <p>23 reports you use the phrase specific causation. What do</p> <p>24 you understand the words specific causation to mean?</p> <p>25 MR. SCHMIDT: Objection, form.</p>	<p>68</p> <p>1 carefully in my report, you know, when you're doing</p> <p>2 scientific studies, you can, for instance, create</p> <p>3 identical materials and randomly assign people to</p> <p>4 different conditions and -- and see whether, you know, a</p> <p>5 candidate who's labeled as female versus a candidate</p> <p>6 who's labeled as male who has otherwise absolutely</p> <p>7 identical information about them is, on average, treated</p> <p>8 differently. That's how we typically really causally</p> <p>9 scientifically nail down whether discrimination occurs</p> <p>10 or not.</p> <p>11 And even in those circumstances, we can't</p> <p>12 for sure identify specific individuals as having</p> <p>13 discriminated or not because they're randomly assigned</p> <p>14 to different conditions; but we can say that overall</p> <p>15 there's a tendency for discrimination to occur or not</p> <p>16 occur in these circumstances. So we can't really do</p> <p>17 that in an individual case.</p> <p>18 And, of course, we talked about social</p> <p>19 desirability biases. I don't think if I interviewed</p> <p>20 Dean Wood and said, "Did you intentionally</p> <p>21 discriminate," or "Did you discriminate," that, you</p> <p>22 know -- or if I provided her with identical materials</p> <p>23 from a man versus a woman that she would, you know --</p> <p>24 she wouldn't be likely to discriminate because she</p> <p>25 would -- it would be clear that they were identical.</p>
<p>67</p> <p>1 A. My understanding of that is that in this kind</p> <p>2 of legal framework that we've talked about, about social</p> <p>3 framework testimony, specific causation would refer to</p> <p>4 what -- you know, what -- specific causal conclusions</p> <p>5 about what happened in the case.</p> <p>6 Q. (BY MR. GIBSON) And is it general- -- is it</p> <p>7 possible or feasible to conduct a study to determine</p> <p>8 whether discrimination occurred in a specific case to an</p> <p>9 individual?</p> <p>10 A. After the fact, I think it's really impossible</p> <p>11 to completely scientifically nail that down in the</p> <p>12 context of a legal case; and so I just don't think</p> <p>13 someone who claims they can do that really can do that.</p> <p>14 Q. And is one of the reasons it's very difficult</p> <p>15 to rule out alternative explanations for behavior other</p> <p>16 than discrimination or bias?</p> <p>17 A. I'm sorry. Can you repeat that?</p> <p>18 Q. Sure. Is one of the reasons it's difficult or</p> <p>19 near impossible to do that in a scientific manner is</p> <p>20 because it's difficult to rule out alternative</p> <p>21 explanations for behaviors other than discrimination or</p> <p>22 bias?</p> <p>23 A. Okay. So -- so when you say -- the fact that</p> <p>24 you can't conduct a study is what you're saying, right?</p> <p>25 Okay. So, as I explained, I think, very clearly and</p>	<p>69</p> <p>1 So, you know, you really can't completely nail that</p> <p>2 down; and so that's why -- that's why I don't come to a</p> <p>3 scientific conclusion in that section on the case.</p> <p>4 Q. In the Mullenix report you state the</p> <p>5 following -- I'm just going to quote this -- "For</p> <p>6 specific causation is explained below. Social framework</p> <p>7 experts often do not apply a scientific certainty</p> <p>8 standard because it may not be possible or not feasible</p> <p>9 to conduct a rigorous study to determine whether</p> <p>10 discrimination occurred in a specific case to an</p> <p>11 individual plaintiff. In such cases experts cannot</p> <p>12 testify with scientific certainty about whether</p> <p>13 discrimination occurred."</p> <p>14 A. Right.</p> <p>15 Q. You continue -- do you agree with that</p> <p>16 statement with respect to this case as well?</p> <p>17 A. Yes, I agree with that statement with respect</p> <p>18 to this case.</p> <p>19 Q. Similarly, in Mullenix, you -- in your</p> <p>20 Mullenix report you state, "Although I point out ways in</p> <p>21 which general principles can be applied to the current</p> <p>22 case and opine about where case facts are consistent</p> <p>23 with the possibility of discrimination, I expressly note</p> <p>24 that because alternative explanations cannot be ruled</p> <p>25 out, my case opinions do not carry the weight of</p>

<p>70</p> <p>1 scientific certainty."</p> <p>2 A. Right, yes. And I think I was consistent with</p> <p>3 that in the Nikolova report.</p> <p>4 Q. And you -- just again, you would agree that</p> <p>5 that statement also applies to the Nikolova report?</p> <p>6 A. Yes.</p> <p>7 Q. And, finally, Mullenix at 49 you state,</p> <p>8 "Because alternative explanations offered by defendant</p> <p>9 to explain their actions cannot be ruled out</p> <p>10 scientifically, case decision-makers must ultimately</p> <p>11 decide whether they believe discrimination likely did or</p> <p>12 did not occur." Would you agree with that statement as</p> <p>13 applied to this case as well?</p> <p>14 A. Yes, I would.</p> <p>15 Q. Okay. Just to confirm, accordingly, can you</p> <p>16 determine with a scientific certainty whether</p> <p>17 discrimination occurred in this case?</p> <p>18 A. I think I already stated that; but no, I</p> <p>19 cannot.</p> <p>20 Q. I would like to start -- go to page 10 of your</p> <p>21 report. This is where you provide what I'll call a</p> <p>22 summary of scientific research on stereotyping and</p> <p>23 discrimination. Most of these studies -- are most of</p> <p>24 these studies done in an experimental environment or in</p> <p>25 a real-world environment?</p>	<p>72</p> <p>1 so on.</p> <p>2 Q. And some of the studies that you cite include</p> <p>3 two different kinds of experiments within the same</p> <p>4 article or publication, one of which would be what I'll</p> <p>5 call a Lab A study that's, you know, sort of in what</p> <p>6 I'll call a scenario. It's not the real world, but it's</p> <p>7 sort of a hypothetical scenario where undergraduates are</p> <p>8 asked to do -- respond in some way. And then they also</p> <p>9 include what are called audit studies. What's an audit</p> <p>10 study?</p> <p>11 MR. SCHMIDT: Objection, form.</p> <p>12 A. An audit study would be to send out, for</p> <p>13 example, resumes to employers to a -- you know, it's</p> <p>14 often a large number of employers; and then vary,</p> <p>15 through random assignment, whether you, for instance,</p> <p>16 send the identical resume with a male name versus a</p> <p>17 female name to see on some measure, like, what</p> <p>18 percentage of people get called back or, you know, those</p> <p>19 kinds of measures. It's a -- it's an attempt to test</p> <p>20 those things out in more of a, you know, business</p> <p>21 setting.</p> <p>22 Q. (BY MR. GIBSON) In addition, some of the</p> <p>23 studies that you cite also look at actual data from</p> <p>24 real-world situations, including universities, such as</p> <p>25 student teacher evaluation data?</p>
<p>71</p> <p>1 A. There's a mix.</p> <p>2 Q. What are -- what would you say most of them</p> <p>3 are conduct- -- how most of them are conducted?</p> <p>4 A. I would have to go through and count.</p> <p>5 Q. Okay.</p> <p>6 A. So I wasn't keeping count. I -- you know, so</p> <p>7 I couldn't tell you right now.</p> <p>8 Q. On page 6 of your report, if I recall</p> <p>9 correctly, you say -- let's see. Where did I see this?</p> <p>10 The second paragraph, about two-thirds of</p> <p>11 the way down, you say, "Social scientific research on</p> <p>12 stereotyping and discrimination normally involves</p> <p>13 voluntary participants who receive assurances (which</p> <p>14 they can rely on) that their responses are anonymous and</p> <p>15 confidential."</p> <p>16 A. Yes. Certainly in the context of, say, a</p> <p>17 laboratory experiment, yeah.</p> <p>18 Q. And are these often -- experiments often</p> <p>19 conducted using undergraduate students at universities</p> <p>20 and colleges?</p> <p>21 A. They're certainly often conducted that way,</p> <p>22 but they're also often conducted in more real-world</p> <p>23 settings or with people who have -- increasingly with</p> <p>24 samples that are obtained on the internet for getting</p> <p>25 more diverse samples, more people who are employed, and</p>	<p>73</p> <p>1 A. Sure. Some of them look at real-world</p> <p>2 situations and they're not -- you know, sometimes</p> <p>3 they're not necessarily experimental studies, that is,</p> <p>4 not necessarily randomly assigned and so on. And then</p> <p>5 there's also studies that are in organizations, for</p> <p>6 instance, looking at rates at which men and women</p> <p>7 receive organizational rewards as compared to how</p> <p>8 they're evaluated in performance.</p> <p>9 Q. Did you do any of those kind of studies here</p> <p>10 in this case with respect to data from the University of</p> <p>11 Texas, for instance?</p> <p>12 A. Again, I think I've already answered that</p> <p>13 question. I did not do that kind of study at the</p> <p>14 University of Texas. I think that would be a</p> <p>15 prohibitive burden on the Plaintiff to perform that kind</p> <p>16 of study, and it wouldn't ultimately resolve that</p> <p>17 question of saying to a scientific certainty that the</p> <p>18 Plaintiff herself was discriminated against.</p> <p>19 So I think you're -- you're setting -- if</p> <p>20 you're suggesting that that would be something that a</p> <p>21 plaintiff should do, I think for an individual</p> <p>22 plaintiff, that's -- that's -- that's asking -- that's</p> <p>23 really placing undue burden on them from my perspective;</p> <p>24 and, also, it's --</p> <p>25 Q. Is that your personal --</p>

<p style="text-align: right;">74</p> <p>1 A. -- not ultimately going to resolve the issue</p> <p>2 from a scientific perspective.</p> <p>3 Q. Is that your personal opinion, or are you</p> <p>4 offering that as an expert that that would provide an</p> <p>5 undue burden on a plaintiff?</p> <p>6 A. Well, as you -- as you noted, everything's</p> <p>7 informed by my experience as -- you know, as a social</p> <p>8 scientist. So I'm -- I'm not giving you a legal</p> <p>9 opinion. That's certainly -- "undue burden," I know,</p> <p>10 has legal resonance. That would be for a judge to</p> <p>11 decide.</p> <p>12 My argument would be that: Look, to</p> <p>13 demand that an individual plaintiff fund such a large-</p> <p>14 scale study that isn't ultimately -- I mean, that could</p> <p>15 establish that there's a pattern at the university;</p> <p>16 but it ultimately is not going to establish with a</p> <p>17 scientific certainty that this individual plaintiff was</p> <p>18 discriminated against.</p> <p>19 So I think I'm just pointing out that</p> <p>20 there's kind of a disconnect here between that and then</p> <p>21 saying with a scientific certainty that this particular</p> <p>22 plaintiff was discriminated against.</p> <p>23 Q. So wouldn't that evidence be more relevant</p> <p>24 than evidence, say, about student teacher evaluations</p> <p>25 from universities in the Netherlands?</p>	<p style="text-align: right;">76</p> <p>1 to be credible and true, that provides a background --</p> <p>2 background information."</p> <p>3 Q. Again, did you ask for that information?</p> <p>4 A. I did not ask for that information.</p> <p>5 Q. Okay. I'd like to ask you about some of the</p> <p>6 studies that you cite in your report. First, if you</p> <p>7 could turn to page 11 of your report, in the first</p> <p>8 paragraph, third line, you say, "Specifically, in 6</p> <p>9 studies involving over 4,000 individuals that included</p> <p>10 both supervisors' performance evaluations and ratings of</p> <p>11 subordinates 'promotion potential,' even though women</p> <p>12 were generally rated as performing better than men,</p> <p>13 supervisors were more likely to rate men as having</p> <p>14 higher promotion potential than women." Do you see</p> <p>15 that?</p> <p>16 A. Yes, I see that.</p> <p>17 Q. Okay. In Footnote 20 you cite the Roth</p> <p>18 publication, "A meta-analysis of gender group</p> <p>19 differences for measures of job performance"?</p> <p>20 A. Correct, yes.</p> <p>21 Q. Great. Let's talk about the Roth study. It's</p> <p>22 going to take me just a second to get that to you. Hold</p> <p>23 on just a second.</p> <p>24 I have to relabel things to make sure I</p> <p>25 give them the proper exhibit numbers. Just a second.</p>
<p style="text-align: right;">75</p> <p>1 A. Well, the research that I cite isn't confined</p> <p>2 to the Netherlands, for one thing.</p> <p>3 Q. I didn't say it was. I didn't say it was.</p> <p>4 That wasn't my question. So if you could answer the</p> <p>5 question instead of taking every question in a</p> <p>6 defensive tone, we could just get through this faster.</p> <p>7 I wasn't -- I didn't say all of your studies were from</p> <p>8 the Netherlands. I'm asking -- I asked you a question.</p> <p>9 A. I'm not trying to be defensive. I'm just</p> <p>10 trying to -- to make sure that I'm not --</p> <p>11 Q. You made a statement about the studies that</p> <p>12 are cited in your report that have -- that is not what I</p> <p>13 said in my question.</p> <p>14 A. If you're asking would it be useful to have</p> <p>15 data on how male and female professors are evaluated</p> <p>16 within the Cockrell School, for instance, that would be</p> <p>17 useful data.</p> <p>18 Q. Did you ever request that data?</p> <p>19 A. I did not request that data.</p> <p>20 Q. Did you ever request data about the history</p> <p>21 of promotion and tenure decisions within the Cockrell</p> <p>22 School of Engineering or UT Austin in general?</p> <p>23 A. Well, there was, you know, information about</p> <p>24 that in the materials I received; and I -- and I cited</p> <p>25 those and said, "Well, if -- you know, if that's found</p>	<p style="text-align: right;">77</p> <p>1 (Exhibit 2 marked.)</p> <p>2 Q. (BY MR. GIBSON) Okay. It should be in the</p> <p>3 chat function now.</p> <p>4 THE WITNESS: Hey, am I still with you</p> <p>5 guys --</p> <p>6 MR. GIBSON: Yes.</p> <p>7 THE WITNESS: -- because my Zoom is</p> <p>8 saying it's not responding? All right. I think it's</p> <p>9 just --</p> <p>10 MR. GIBSON: Yes, we can still hear you.</p> <p>11 THE WITNESS: Okay. All right. I think</p> <p>12 it was just having trouble doing both, opening the</p> <p>13 document and -- okay. Let's see if that will open.</p> <p>14 A. Yeah, I see it.</p> <p>15 Q. (BY MR. GIBSON) Great. Does this appear to</p> <p>16 be the study that you're referring to in Footnote 20?</p> <p>17 A. It does appear to be that study.</p> <p>18 Q. Great. Do you know the timeframe of the</p> <p>19 studies that were cited -- well, first, what's a</p> <p>20 meta-analysis?</p> <p>21 A. Okay. A meta-analysis is a statistical</p> <p>22 conglomeration of a lot of previous data. So you try</p> <p>23 to identify, you know, out there in the research</p> <p>24 literature among published studies and even potentially</p> <p>25 among what seem to be high-quality non-published</p>

<p>78</p> <p>1 studies, like in a dissertation, the kind of universe of</p> <p>2 studies that might be relevant to a specific question</p> <p>3 and then try to aggregate them together statistically to</p> <p>4 try to estimate -- you know, to try to test a question</p> <p>5 or multiple questions.</p> <p>6 Q. Okay. Do you know the timeframe of the study</p> <p>7 used in this report?</p> <p>8 A. Let's see. I can't offhand tell you the</p> <p>9 timeframe. I think it took -- when was the study</p> <p>10 published? The study -- the data are, I'm sure, older</p> <p>11 than exactly when the study was published. It was</p> <p>12 published in 2016; but the data, I believe, were from</p> <p>13 decades of research. That's not uncommon with a</p> <p>14 meta-analysis.</p> <p>15 Q. Okay. So do you think you can make -- so</p> <p>16 would it surprise you to find out that some of the</p> <p>17 studies cited in here are from the 1960s?</p> <p>18 A. I wouldn't be shocked, no.</p> <p>19 Q. Do you think that that affects the application</p> <p>20 of this meta-analysis to current, present-day work</p> <p>21 environments?</p> <p>22 A. Well, I'm trying to remember if in this</p> <p>23 meta-analysis they looked at time of publication as a</p> <p>24 variable because that would be kind of a typical thing</p> <p>25 to do to see if the tendency had diminished or not; but</p>	<p>80</p> <p>1 more likely. So that would be something that would be a</p> <p>2 potential moderating factor to look at in -- in a</p> <p>3 meta-analysis and to see if that -- you know, that shows</p> <p>4 up in the -- in the study results --</p> <p>5 Q. Did you --</p> <p>6 A. -- if they even categorize that.</p> <p>7 Q. Did you determine whether or not any of those</p> <p>8 additional analyses had been done in this study before</p> <p>9 you relied upon it and using it in this case?</p> <p>10 A. Okay. It's one of many studies that I cited.</p> <p>11 I, again, can't -- don't have encyclopedic recall for</p> <p>12 details of everything that I cited. You know, this is</p> <p>13 a meta-analysis published in a peer-reviewed good</p> <p>14 journal. So I -- I -- you know, I can't -- I reviewed</p> <p>15 it. I thought it was worth including. I can't tell</p> <p>16 you everything about the study in a -- kind of like you</p> <p>17 pick something out of a specific page, whatever, I'm</p> <p>18 happy to look at it; but I can't recall all the details</p> <p>19 at this -- offhand, at this moment.</p> <p>20 Q. You rely on this study in reaching the</p> <p>21 conclusions in this -- your opinions in this case; is</p> <p>22 that correct?</p> <p>23 A. I rely on this as part of the framework, the</p> <p>24 scientific framework, and not just this study. You'll</p> <p>25 see that immediately following, there is a study that,</p>
<p>79</p> <p>1 I'd have to go back because I'm not an encyclopedia. I</p> <p>2 can't remember for sure if they used the time of</p> <p>3 publication as a moderating factor.</p> <p>4 Q. Okay. Do you know -- and do you think it's</p> <p>5 appropriate to use unpublished research in a</p> <p>6 meta-analysis?</p> <p>7 A. Well, usually, the people who are conducting</p> <p>8 the meta-analysis, which goes through peer review, the</p> <p>9 bulk of the studies are usually published; but they also</p> <p>10 try to -- to get unpublished studies because, you know,</p> <p>11 of concern that they're getting a really good sample.</p> <p>12 And they try to, you know, look at indicators of quality</p> <p>13 of the study. So -- so that's not unusual at all for --</p> <p>14 for authors to try to do that, but they usually have</p> <p>15 criteria that they will impose to try to weed out</p> <p>16 studies that they see as low in quality.</p> <p>17 Q. Do you -- do you think it's appropriate to --</p> <p>18 do you think it matters what the context of the study</p> <p>19 was in terms of the specific work environment that was</p> <p>20 being examined?</p> <p>21 A. Well, again, that could be a moderating</p> <p>22 variable that people might look at. It really depends.</p> <p>23 It really depends, you know. For instance, as you</p> <p>24 mentioned, we -- the research -- research tends to find</p> <p>25 that in male-dominated occupations, discrimination is</p>	<p>81</p> <p>1 arguably, is much better, larger, looking at actual</p> <p>2 allocation of rewards rather than ratings of</p> <p>3 promotability, a much larger sample, you know, right?</p> <p>4 And -- and so when you -- you also look for whether</p> <p>5 there's -- you know, these are both meta-analyses.</p> <p>6 Those -- those are useful studies to cite.</p> <p>7 Q. In your statement about the Roth article, you</p> <p>8 say that there's -- "in 6 studies involving over 4,000</p> <p>9 individuals." Can -- I don't understand what studies</p> <p>10 you're citing to because in this meta-analysis, they</p> <p>11 look at much more than six studies.</p> <p>12 A. Right, right, they did.</p> <p>13 Q. Can you --</p> <p>14 A. But only in six studies did they have both</p> <p>15 performance evaluations and ratings of promotability.</p> <p>16 So the number of studies in which they had ratings of</p> <p>17 promotability was much lower than their overall sample.</p> <p>18 So I didn't want to claim that in their overall sample</p> <p>19 that's what they found because they needed to have those</p> <p>20 measures included in the study. This is part of the --</p> <p>21 you know, the -- when you're doing a meta-analysis,</p> <p>22 different studies look at different measures. And so,</p> <p>23 you know, this is -- this is really very common that</p> <p>24 there might be a subset of studies where you could test</p> <p>25 a question that aren't true for the bulk of studies.</p>

<p>94</p> <p>1 I will point out that that's quoting</p> <p>2 directly from the article which was accepted in what's</p> <p>3 considered a relatively prestigious journal and subject</p> <p>4 to very significant peer review that is in the applied</p> <p>5 area of really trying to study organizations.</p> <p>6 Q. Will you turn to page 1530 of the article,</p> <p>7 please?</p> <p>8 A. Okay. 15...</p> <p>9 Q. And I'll pull this up just so we're all</p> <p>10 looking at the same thing.</p> <p>11 A. Yeah.</p> <p>12 Q. If I can do this appropriately.</p> <p>13 Let me do this. Okay. We are now</p> <p>14 looking at the Joshi article. Is that what you see on</p> <p>15 your screen, Dr. Glick?</p> <p>16 A. Well, I'm looking at page 1530 and -- yeah.</p> <p>17 So, yeah, that article -- but I'm -- I was looking --</p> <p>18 oh, shoot. Now, where...</p> <p>19 Q. That's okay. I'll flip to it. We can all</p> <p>20 look at the same thing. I apologize --</p> <p>21 (Simultaneous speakers.)</p> <p>22 A. The -- the font is really small on your</p> <p>23 screen, and now I sort of lost --</p> <p>24 Q. Let me restart --</p> <p>25 A. That's better.</p>	<p>96</p> <p>1 the null effect, which would be the lightest, white-</p> <p>2 colored triangle in the middle, which would mean there</p> <p>3 were no statistic -- statistical significance between</p> <p>4 the evaluations of men and women; but in other studies,</p> <p>5 there was quite a bit of difference between the</p> <p>6 evaluations of men and women, correct?</p> <p>7 A. Yes.</p> <p>8 Q. And do you know the specific context of those</p> <p>9 data -- and it's okay if the answer's no. I just want</p> <p>10 to confirm: Do you know the specific context of the</p> <p>11 data points that are -- show a statistical difference</p> <p>12 that favor women?</p> <p>13 A. From this chart, no.</p> <p>14 Q. Great. Okay. And if you go to the next</p> <p>15 chart, that is the chart of organizational rewards where</p> <p>16 most but not all of the data points of the studies show</p> <p>17 a -- that men are receiving more organizational rewards</p> <p>18 than women because that -- those are the data points</p> <p>19 that are on the right side of the chart; is that</p> <p>20 correct?</p> <p>21 A. That is correct.</p> <p>22 Q. However, there are also data points on the</p> <p>23 left side of the chart; is that correct?</p> <p>24 A. Yes.</p> <p>25 Q. And those would be studies that show a</p>
<p>95</p> <p>1 Q. Let me restart --</p> <p>2 A. That's better.</p> <p>3 Q. Dr. Glick, can I restart the questioning?</p> <p>4 A. Sure.</p> <p>5 Q. Does Exhibit 3 appear to be the Joshi article</p> <p>6 that's cited in Footnote 21 of your report?</p> <p>7 A. Yes, it does.</p> <p>8 Q. Great. I'm going to go to -- okay. I am on</p> <p>9 page 1530 of Exhibit 3, Figure 1. Do you see that?</p> <p>10 A. Yes, I do.</p> <p>11 Q. Great. And does this show the various</p> <p>12 studies -- is it your understanding that each data point</p> <p>13 is a study that was used in the meta-analysis?</p> <p>14 A. I would like to re-read this section just to</p> <p>15 make sure; but that's what it appears to be, yes.</p> <p>16 Q. Okay. And does that indicate in some studies</p> <p>17 the effect was in favor of women, and in some studies</p> <p>18 the effect -- let's talk about performance evaluations</p> <p>19 first.</p> <p>20 A. Uh-huh.</p> <p>21 Q. In some studies the effect of the performance</p> <p>22 evaluations favored women; in some studies, it favored</p> <p>23 men?</p> <p>24 A. Correct.</p> <p>25 Q. Right. And that in some studies it was within</p>	<p>97</p> <p>1 statistically significant benefit where women receive</p> <p>2 more of the organizational rewards than men; is that</p> <p>3 correct?</p> <p>4 A. Correct.</p> <p>5 Q. Do you know the context of the specific</p> <p>6 studies that are on the left side of the -- this</p> <p>7 particular chart?</p> <p>8 A. Not from this chart.</p> <p>9 Q. Okay. Are you aware that this meta-analysis</p> <p>10 looked at studies from all over the world?</p> <p>11 A. I, again, want to review the paper before</p> <p>12 making absolute statements about that.</p> <p>13 Q. Okay. Well, let's do that. Let's go to page</p> <p>14 1536, and I'm going to highlight a section. Do you see</p> <p>15 the -- do you understand that the asterisk -- well, I'm</p> <p>16 on page 1536, which is a list of references; and at the</p> <p>17 beginning of the list of references it says, "References</p> <p>18 marked with an asterisk indicate studies included in the</p> <p>19 meta-analysis." Did I read that correctly?</p> <p>20 A. Yes.</p> <p>21 Q. So --</p> <p>22 (Simultaneous speakers.)</p> <p>23 Q. -- that I'm highlighting that starts with</p> <p>24 Bhatnagar, B-H-A-T-N-A-G-A-R --</p> <p>25 A. Yes.</p>

<p>98</p> <p>1 Q. -- does that appear to be a study that was 2 included in the meta-analysis?</p> <p>3 A. That appears to be a study included in the 4 meta-analysis.</p> <p>5 Q. Does that study appear to relate to Indian 6 knowledge workers?</p> <p>7 A. Correct.</p> <p>8 Q. That's -- it's your understanding that's in 9 India?</p> <p>10 A. Yes, that was -- that's the implication, yes.</p> <p>11 Q. Do you know if that study -- do you know where 12 that study fell on the charts re- -- that are really 13 the -- as I understand, those two charts summarize the 14 data that you really highlight from this study about the 15 differences between those two effects, correct?</p> <p>16 A. Right.</p> <p>17 Q. Do you know where that chart appears on that 18 study?</p> <p>19 A. You mean --</p> <p>20 Q. I mean, where this study appears on the chart. 21 My apologies.</p> <p>22 A. Right. Okay. I figured that's what you were 23 asking. No, I do not.</p> <p>24 Q. And if I told you that there's numerous 25 studies in here from China, would that surprise you?</p>	<p>100</p> <p>1 Q. Okay. So are you aware that there are studies 2 in here about Swiss bankers?</p> <p>3 A. I would trust you if you told me that there 4 were.</p> <p>5 Q. Great. Going back to the chart, I think there 6 may be one or two studies in here about university 7 faculty members.</p> <p>8 A. There may be.</p> <p>9 Q. Do you know if the studies about university -- 10 looking at the charts on page a 1530 of this study, do 11 you know if the studies about university faculty members 12 are actually the studies that show that there's a 13 preference in favor of women --</p> <p>14 A. I can't tell that from this chart.</p> <p>15 Q. -- in the university faculty context?</p> <p>16 Sorry. In the university faculty 17 context?</p> <p>18 A. Sorry. I thought you were finished with your 19 question. I'm looking at the paper and not your face, 20 so it's hard to tell when you're done.</p> <p>21 You can't tell from this chart, no.</p> <p>22 Q. So it's possible --</p> <p>23 A. You can't tell either way.</p> <p>24 Q. So it's possible that those studies that are 25 about universities or faculty members actually show the</p>
<p>99</p> <p>1 A. It would not shock me, no.</p> <p>2 Q. Okay. Do you think it's appropriate to make 3 conclusions about whether or not something is -- well, 4 do you think it's appropriate to rely on studies from 5 China in making conclusions about how social science 6 research may be applicable to the Nikolova case?</p> <p>7 A. It really depends on what that study is about.</p> <p>8 Q. Great. I'm asking: Do you have an answer 9 with respect to the specific studies used in this 10 meta-analysis from China? Was it -- is it appropriate 11 to use those studies to make conclusions about the work 12 environment at the University of Texas at Austin in 2021 13 that Dr. Nikolova was subject to?</p> <p>14 A. Well, again, I think that -- that it would be 15 better to have studies that are closer to the context or 16 certainly that are, you know, closer to the setting; but 17 that doesn't mean that these studies are useless. And I 18 think if you go back to the chart, you see that -- I 19 mean, it's a pretty striking chart -- that there are 20 relatively few studies that are showing an advantage to 21 women. And if you look at that chart, I mean, there are 22 studies in here that are from -- from the U.S. or 23 northern Europe that might be culturally somewhat 24 similar in terms of overall gender equality. So I think 25 that the meta-analysis is still highly relevant.</p>	<p>101</p> <p>1 opposite of what the meta-analysis shows?</p> <p>2 A. That could be possible, yes.</p> <p>3 Q. And just to confirm, have you done any 4 separate analysis of studies involving faculty members 5 at univers- -- at American universities to determine 6 what those studies show in this type of context?</p> <p>7 A. Well, I believe in my report there are studies 8 that are about women in engineering that -- you know, 9 women in engineering in the U.S., organizational 10 climate, those sorts of things. There are certainly 11 reports on that that I cite.</p> <p>12 Q. Okay. But in terms of -- other than the 13 reports you cite from -- I think most of those are from 14 other social scientists. Have you done any analysis of 15 the studies that specifically relate to faculty members 16 in the United States?</p> <p>17 MR. SCHMIDT: Objection, form.</p> <p>18 A. I'm not sure what you're asking me. Have I 19 done analysis of them? Are you asking if I've done 20 research on university faculty directly?</p> <p>21 Q. (BY MR. GIBSON) Yes.</p> <p>22 A. Is that what you're asking?</p> <p>23 Q. Yes.</p> <p>24 A. So have I researched that, like, done a study 25 myself and researched --</p>

<p>106</p> <p>1 Q. (BY MR. GIBSON) Again, did you conduct an 2 assessment of the studies used in this meta-analysis 3 before applying the conclusions of the meta-analysis to 4 this case? 5 A. I'm having trouble understanding what you mean 6 by "conduct an assessment." 7 Q. Did you look at the studies, whether as 8 they're summarized in meta-analysis or individually, 9 separately, to determine whether or not they were 10 applicable to this case, the Nikolova case? 11 A. I looked at the meta-analysis as a whole and 12 the things discussed in the meta-analysis. If you're 13 asking did I go and then read all of the individual 14 studies they cite, no, I did not. 15 Q. Did you even look at the summary of the 16 meta-analysis that -- I'm sorry. Did you even look at 17 the summary of the studies that's in the meta-analysis 18 to determine whether or not those studies were 19 appropriate bases to create conclusions about this case? 20 A. Summary of the -- I don't know what you mean 21 by "summary of the studies." The entire meta-analysis 22 is a summary of the studies. 23 Q. So this -- this meta-analysis -- meta-analysis 24 includes detailed descriptions of all the studies. And 25 my question is: Did you go and review those before</p>	<p>108</p> <p>1 United Kingdom, the Netherlands, Sweden, Israel, 2 Malaysia, Israel again. So you would have been aware 3 that those studies were done in those countries if -- if 4 you would have looked at it as you think you did? 5 MR. SCHMIDT: Objection, form. 6 A. Yes, I would have been aware that there were 7 studies from different parts of the globe and that 8 most of the studies, as you said, were -- were from 9 North America, Europe, where there's considerable 10 cultural similarities when it comes to, you know, bias 11 against women. 12 Q (BY MR. GIBSON) Sure. Your section on -- 13 well, throughout your report you rely on a lot of 14 studies that are certainly more than a decade old and in 15 some cases over 20 years old. Do you have any concerns 16 about relying on studies that are two decades old when 17 making statements about how stereotypes about women in 18 the workplace -- you know, whether such stereotypes 19 exist and how they impact evaluation of women in the 20 workplace? 21 A. It depends. 22 Q. Well, okay. Let's go to Footnote 23. You 23 cite an article by Cohen from 1998 to support the 24 statement, "Other factors exacerbate the likelihood that 25 stereotyping and discrimination will occur. When women</p>
<p>107</p> <p>1 determining that this was an appropriate meta-analysis 2 to use in this case? 3 A. I do not recall. I'd have to look at the 4 study again in detail to answer that question. 5 Q. Let's look at it. If you look at the end, 6 they first list all the studies. That starts on 7 page 142, with a similar use of the asterisk to 8 determine which studies were used in the analysis. 9 There's then a summary of studies by job titles and 10 study year, going back to 1974. There's then an 11 assessment of -- a table of the studies that includes 12 all sorts of information about the studies, where it 13 was -- what country it was from, the study design, the 14 type of participants, the various data, and sex 15 distribution of the studies. Did you look at any of 16 this before determining whether or not this 17 meta-analysis was appropriate to use in the Nikolova 18 report? 19 A. I believe I looked at it. I can't recall how 20 much time I spent on these particular tables. 21 Q. Okay. So if you looked at it, you would 22 have been aware, for instance, that many of these 23 studies come from around the globe. So just looking at 24 page 151, for instance, while most are from the 25 United States, we do see Canada, Austria, Germany, the</p>	<p>109</p> <p>1 or minority group members are underrepresented, they 2 tend to be viewed as a 'poorer fit' for a job than men 3 or majority group members," citing an article from 1998. 4 A. Okay. 5 Q. Do you have any concerns about making 6 statements about the state of the workplace for women in 7 2021 based on research that was published in 1998? 8 A. Well, generally, I rely on my understanding of 9 the research area as a whole. I'm not citing every 10 study that might demonstrate this point. I recently 11 wrote a second edition of a book, co-wrote with my 12 colleague Lori Rudman, on the social psychology of 13 gender, where we reviewed studies in the workplace. 14 That's just recently been published. And these 15 principles have shown generally to hold up pretty well. 16 So this is a case where maybe it would 17 have been smart of me to update the reference; but based 18 on the review that we've conducted, I'm pretty confident 19 in these statements. 20 Q. So are you familiar with the federal rule 21 requiring you to cite the bases upon which your 22 opinions -- the sources upon which your opinions are 23 based? 24 A. I am citing -- 25 MR. SCHMIDT: Objection, form.</p>

<p>110</p> <p>1 Sorry go ahead.</p> <p>2 A. I'm not -- I couldn't cite you the particular</p> <p>3 rule.</p> <p>4 Q. (BY MR. GIBSON) And, similarly, in</p> <p>5 Footnote 32, you cite stereotype -- you cite an article</p> <p>6 from 2002 about gender stereotypes and how that affects</p> <p>7 women and men in high-power, dominance-oriented traits.</p> <p>8 This is on page 13. And do you agree that it would be</p> <p>9 better to have an article that was closer in time than</p> <p>10 1992 -- I'm sorry -- than 2002 to --</p> <p>11 A. Yes, and I do. If you look right below it, I</p> <p>12 cite Rudman, Moss-Racusin, Glick, and Phelan -- the</p> <p>13 Glick is me --</p> <p>14 Q. Uh-huh.</p> <p>15 A. -- in a 2012 review article that reports more</p> <p>16 recent data.</p> <p>17 Q. Uh-huh.</p> <p>18 A. And the same results come out. Gender</p> <p>19 stereotypes show a great deal of stability on some</p> <p>20 dimensions over time.</p> <p>21 Q. But not all dimensions; is that correct?</p> <p>22 A. Correct.</p> <p>23 Q. What dimensions do they not show stability in?</p> <p>24 A. Well, the one dimension where they have shown</p> <p>25 a change -- well, first off, one dimension where they've</p>	<p>112</p> <p>1 and -- and look at that. There's -- there's a, you</p> <p>2 know, recent paper by Alice Eagly that looked at this</p> <p>3 over decades; and I'm trying to recall offhand -- I</p> <p>4 mean, the overall message was that men were still</p> <p>5 favored on agency. I can't remember whether that gap</p> <p>6 was closed significantly or not.</p> <p>7 Q. Well, let's look at a study that you -- if you</p> <p>8 go to Footnote 53 you say, "Stereotypes set up different</p> <p>9 expect-" --</p> <p>10 MR. SCHMIDT: Can you show the document</p> <p>11 that you're referring -- reading from?</p> <p>12 MR. GIBSON: Absolutely. We're back to</p> <p>13 Exhibit Number 1, his report; and I'll pull that up.</p> <p>14 Just a second.</p> <p>15 Q. (BY MR. GIBSON) Hopefully you should be able</p> <p>16 to see that now.</p> <p>17 A. I'm not -- I'm still seeing the old one, I</p> <p>18 think.</p> <p>19 Q. You should see your report starting, Gendered</p> <p>20 Double Standards, at page 18.</p> <p>21 A. Oh, I have my report, so that's fine; but my</p> <p>22 screen has not changed.</p> <p>23 MR. GIBSON: Bob, is that what you're</p> <p>24 seeing?</p> <p>25 MR. SCHMIDT: Yeah, my screen has not</p>
<p>111</p> <p>1 shown a change is over time to actually ascribe more</p> <p>2 communality or warmth traits to women. Another place</p> <p>3 where they've shown a change is to -- to attribute more</p> <p>4 overall competence to women than was previously the</p> <p>5 case.</p> <p>6 Q. In addition, the -- the gender stereotypes</p> <p>7 regarding leadership have been decreasing over time as</p> <p>8 well; is that correct?</p> <p>9 A. I don't know what exactly you're asking.</p> <p>10 Q. So you cite to articles to support the</p> <p>11 argument that leadership qualities are seen as gender</p> <p>12 and the characteristics of good leaders are seen as</p> <p>13 gender.</p> <p>14 MR. SCHMIDT: Objection, form.</p> <p>15 A. Yes, and on agentic characteristics the</p> <p>16 gender -- the stereotypes still remain favoring men on</p> <p>17 these characteristics such as decisive, aggressive --</p> <p>18 aggressive, forceful, dominating, those sorts of things</p> <p>19 which are associated with leadership. So there's still</p> <p>20 this association of stereotypically masculine</p> <p>21 characteristics with leadership.</p> <p>22 Q. (BY MR. GIBSON) But those are changing,</p> <p>23 correct? Those are becoming more androgynous, as the</p> <p>24 research describes?</p> <p>25 A. To some extent, they -- I'd have to go back</p>	<p>113</p> <p>1 changed, either. I'm pulling up the report on my -- on</p> <p>2 my -- as an exhibit.</p> <p>3 MR. GIBSON: I'm not sure why that is.</p> <p>4 MR. SCHMIDT: What page are you referring</p> <p>5 to, page 18?</p> <p>6 MR. GIBSON: Page 18.</p> <p>7 MR. SCHMIDT: Okay. Just give us -- give</p> <p>8 me a second. Okay.</p> <p>9 Q. (BY MR. GIBSON) There we go. Okay. So you</p> <p>10 say that, "Stereotypes set up different expectations for</p> <p>11 men and women, leading to double standards in how people</p> <p>12 evaluate individual women and men. High-status roles</p> <p>13 are typically associated both with stereotypically</p> <p>14 masculine traits (ambition, assertiveness, decisiveness)</p> <p>15 and with men as 'role of [sic] incumbents.'"</p> <p>16 A. Role incumbents.</p> <p>17 Q. Yeah, role incumbents.</p> <p>18 For the first statement you had cited a</p> <p>19 report -- article from 2011, another meta-analysis,</p> <p>20 cited by Koenig; is that correct?</p> <p>21 A. Correct.</p> <p>22 Q. Right. Let's look at that. Hang on just a</p> <p>23 second. I'm trying to make sure I'm not showing you</p> <p>24 guys the wrong thing here.</p> <p>25 (Exhibit 5 marked.)</p>

<p>114</p> <p>1 Q. (BY MR. GIBSON) Okay. Here's the Koenig 2 article. I'm dropping it into the chat, also opening it 3 and showing it on the screen. You should now see the 4 Koenig article. Take a second to look at that. Does 5 this appear to be the article cited in Footnote -- which 6 was is -- 7 A. Yes, it seems to be the article. 8 Q. -- 53? 9 Great. Okay. Are you aware that this 10 article also uses international studies from China, 11 Japan, Egypt, Saudi Arabia, and Turkey? 12 A. I trust that that could well be the case. 13 Again, I'd have to review the article; but the typical 14 procedure in meta-analysis is to try to be inclusive of 15 all published studies. 16 Q. Do you think it's appropriate to use an 17 analysis that relies on, among other things, studies 18 from Saudi Arabia about leadership stereotypes being 19 masculine to make determinations about the way that 20 stereotypes and bias function at the University of Texas 21 at Austin in 2021? 22 A. There are also studies conducted in the 23 United States. That's where a lot of the research tends 24 to be done, so this is not -- this is not a phenomenon 25 that's only demonstrated in Saudi Arabia.</p>	<p>116</p> <p>1 Q. Do you know how the results of that study 2 might be different if, let's say, Saudi Arabia wasn't 3 included? 4 A. Well, again, it's pretty typical in 5 meta-analyses -- and I'd have to go back and look at 6 this particular one to see -- but it's pretty typical to 7 separate some of these issues out to test moderation, 8 that is, is it different, the effects stronger or weaker 9 depending on timing of the study, depending on where 10 it's conducted, those sorts of things. 11 Q. Okay. And this also includes studies from the 12 1970s, going back as far as 1973? 13 A. Yes. Meta-analyses typically include older, 14 as well as more recent, studies; and they also typically 15 look at timing of publication as a possible moderator. 16 Q. Did you look to see if this study did that or 17 whether -- what the potential impact of that was on this 18 study? 19 A. At this time I cannot recall specifically. 20 Q. Turn to page 634 of this study, in particular 21 a section called A Priori Moderators. Is this one of 22 the instances that you're discussing where they look at 23 the year of publication to determine whether or not the 24 results of the study change over time? 25 A. That looks to be the case, yes.</p>
<p>115</p> <p>1 Q. But you -- you haven't -- for this particular 2 point, you didn't cite studies that were done in the 3 United States; instead, you cited a meta-analysis that 4 relies on studies from the United States as well as 5 China, Turkey, Egypt, and Saudi Arabia? 6 A. Yes -- 7 Q. My question is -- 8 A. -- right. 9 Q. And you didn't answer the question. The 10 question is: Do you think it's appropriate to make 11 determinations about how -- what -- how stereotypes and 12 bias function at the University of Texas at Austin in 13 2021 based on studies from China, Saudi Arabia, Egypt, 14 and Turkey? 15 MR. SCHMIDT: Objection, form. 16 A. If that were based solely on studies from 17 those countries, the answer would be "no." But my 18 opinion is not based solely on studies from those 19 countries. 20 Q. (BY MR. GIBSON) That statement is based 21 solely -- the only support you provide for that 22 particular statement in your report for Footnote 53 is 23 this study, correct? 24 A. Which -- yes, this study which includes data 25 from the U.S. as well as from other countries.</p>	<p>117</p> <p>1 Q. Great. Does it say that, "Evidence of 2 increasing" -- the first sentence -- "Evidence of 3 increasing androgyny of the leadership stereotype over 4 publication years emerged in all three paradigms, 5 including in the multiple regression equations that 6 controlled for moder- -- other moderator variables"? 7 A. Where are you seeing that? 8 Q. The first sentence of A Priori Moderators, 9 Year of publication. 10 A. Right. Okay. I see. It's really very small. 11 Can you make that a little larger? 12 Q. I'm happy to. You should also have it in 13 your -- to be able to -- 14 A. Yeah, it's easier if I look at it here because 15 it's -- 16 Q. I'm happy to do that. 17 A. -- not slowing us down. 18 Q. That's about as big as I can make it. 19 A. No, that's good. That's good. Good, good, 20 good. 21 Okay. So, "Evidence of increasing 22 androgyny of the leadership stereotype over publication 23 years emerged," right, "...including in the multiple 24 regression equations that controlled for all [sic] other 25 moderator variables (albeit as a mod- -- a marginal</p>

<p>118</p> <p>1 effect in the agency-communion multiple meta- 2 regression.)"</p> <p>3 Q. So, in conclusion, that paragraph says, 4 "Thus, our conclusion is that leadership now, more than 5 in the past, appears to incorporate more feminine 6 rational [sic] qualities, such as sensitivity, warmth, 7 and understanding, thus adding them to the masculine 8 dominance and strength qualities traditionally 9 associated with leadership?"</p> <p>10 A. Sure.</p> <p>11 Q. So then this was in 2011, based on -- so this 12 was a decade ago they had already come to this 13 conclusion, correct?</p> <p>14 A. Yes, this conclusion but it doesn't mean that 15 there's the -- that discrimination has disappeared on 16 this dimension.</p> <p>17 Q. Do you --</p> <p>18 A. It means that there has been change over time, 19 yes.</p> <p>20 Q. Do you know whether or not more recent 21 studies -- what recent studies show, within the past 22 decade within the United States?</p> <p>23 A. I'd have to look specifically at this paradigm 24 in the last decade; and offhand, I can't tell you of -- 25 right now I can't -- you know, I can't cite a study for</p>	<p>120</p> <p>1 spoke with Dr. Glick, it took the whole day, I think, 2 so.</p> <p>3 MR. SCHMIDT: Okay.</p> <p>4 MR. GIBSON: I hope not to take the full 5 day again, but I don't -- I would -- if I had to make an 6 estimate, not holding me to it, I would say hopefully 7 we'll be done by 4:00 o'clock.</p> <p>8 MR. SCHMIDT: All right. Sounds great.</p> <p>9 MR. NOTZON: So, Dr. Glick, you can make 10 the call on when you want to break for lunch.</p> <p>11 THE WITNESS: Now's fine.</p> <p>12 MR. GIBSON: Great. We'll take a one- 13 hour lunch break. Be back at 1:15. It's now 12:13. So 14 the time back is 1:15. Thanks everybody.</p> <p>15 MR. SCHMIDT: Thank you.</p> <p>16 THE VIDEOGRAPHER: The time is 12:13 p.m. 17 We are off the record.</p> <p>18 (Off the record from 12:13 to 1:27 p.m.)</p> <p>19 THE VIDEOGRAPHER: The time is 1:27 p.m. 20 We are back on the record.</p> <p>21 Q (BY MR. GIBSON) Dr. Glick, I would like to 22 move forward.</p> <p>23 Oh, first, before we proceed, we just 24 had about a little over an hour break for lunch. Did 25 you want to go back and address, I think it was,</p>
<p>119</p> <p>1 you.</p> <p>2 Q. Okay.</p> <p>3 THE VIDEOGRAPHER: My apologies for 4 interrupting. Mr. Glick, we can hardly see you on 5 camera.</p> <p>6 THE WITNESS: Oh, I'm sorry. I'm sorry. 7 Yeah, I'm looking at the other stuff on the screen. 8 Sorry about that.</p> <p>9 MR. NOTZON: While you're at a break in 10 your questioning right now, are --</p> <p>11 MR. GIBSON: Sure.</p> <p>12 MR. NOTZON: -- do we want to talk about 13 a lunch break; or are you going to try to push through 14 and finish?</p> <p>15 MR. GIBSON: I'm happy to talk about a 16 lunch break. I don't think we'll be able to push 17 through and finish. So I would suggest we take a lunch 18 break. I think now is a good time. I'm happy to take 19 it now. If other people have preferences, I'm happy to 20 follow other people; but we -- we are going to need a 21 lunch break.</p> <p>22 MR. SCHMIDT: And do you have any 23 estimate of how much longer you're going to go this 24 afternoon?</p> <p>25 MR. GIBSON: I mean, the last time I</p>	<p>121</p> <p>1 Footnote 20, the Roth article, regarding the questions I 2 had asked about that?</p> <p>3 A. No.</p> <p>4 Q. Okay. We will proceed.</p> <p>5 Okay. So I'd like to move to the portion 6 of your report regarding the -- that's entitled -- it 7 starts on page 21, entitled Discrimination From Below: 8 Bias In Teaching Evaluations.</p> <p>9 First off, let's talk about Footnote 70.</p> <p>10 That's the Clayson article from 2009. Do you see that?</p> <p>11 A. Yes.</p> <p>12 Q. Great. Do you -- do you know if there's any 13 time limitation on that study?</p> <p>14 A. Again, I'm not sure if you're asking what the 15 timeline is of which -- which publication years they 16 included.</p> <p>17 Q. Do you know if there's any limit on the 18 publication years of that study?</p> <p>19 A. Not that I -- not that I can tell you right 20 now.</p> <p>21 Q. And do you -- are you aware that the study 22 includes -- the meta-analysis includes individual 23 studies that go back as far as 1953?</p> <p>24 A. Again, I'd have to review the study; but I 25 trust that there might be some older -- older studies in</p>

<p style="text-align: right;">122</p> <p>1 there. Again, meta-analyses try to be very inclusive; 2 and then they often look at things like publication 3 year, as we talked about with prior meta-analyses. 4 Q. Do you know if publication year was looked at 5 in this analysis? 6 A. Offhand, I can't tell you. I'd have to go 7 back and review it. 8 Q. Okay. Do you know if the case -- any 9 difference between the types of classes that were being 10 evaluated, like, math versus physics, versus business 11 classes? 12 A. Again, I'd have to go back and review the 13 details of the study of the meta-analysis. 14 Q. Okay. Do you know if the meta-analysis 15 included engin- -- any engineering classes? 16 A. Again, I'd have to go back and look at the 17 details of the meta-analysis. 18 Q. Okay. Did you conduct any sort of analy- -- 19 did you review the details of the meta-analysis before 20 including it as a basis upon which you -- one of the 21 bases of the conclusions in your report in this case? 22 A. I read it if that's what you mean. 23 Q. Did you read it as part of drafting this 24 report or have you just -- when you say you've read it, 25 does that just mean you've read it at some point in the</p>	<p style="text-align: right;">124</p> <p>1 that study was from, meaning juris- -- what country it 2 was from? 3 A. I couldn't tell you for sure, but -- I'm 4 thinking maybe Germany, but I can't remember for sure. 5 Q. Let's pull that up. 6 MR. GIBSON: Okay. Sorry. I had all 7 this ready to go, of course; and then my computer made 8 me reboot. So I had to shut it all down. So sorry it's 9 taking a little longer. 10 (Exhibit 6 marked.) 11 Q. (BY MR. GIBSON) You should have Exhibit 6 in 12 your -- in the chat. I'll also pull it up on the share 13 screen so that you can confirm that Exhibit 6 is the 14 article you're referring to in Footnote 75. 15 A. Yes, that looks like it. 16 Q. Great. Are you aware that this study uses a 17 database from a French university? Do you see that in 18 the very first line? 19 A. I see that. So I got Europe right. 20 Q. Great. And are you aware that this study was 21 based on sort of introductory classes at this French 22 university? 23 A. Again, I'd have to review all of it; but... 24 Q. Sure. Do you -- 25 A. I assume that you're being correct on that.</p>
<p style="text-align: right;">123</p> <p>1 past or you've read it as part of drafting the report in 2 this case? 3 A. I've read it as part of drafting the report. 4 Q. And so that means you made a conclusion that 5 the -- it was an appropriate study upon which to base 6 your opinions in this case? 7 A. Sure. 8 Q. Okay. 9 MR. NOTZON: Just to clarify, it was "an" 10 appropriate, instead of inappropriate, correct? 11 MR. GIBSON: That's correct. 12 (Laughter.) 13 Q. (BY MR. GIBSON) An, A-N, appropriate study. 14 A. One of other appropriate studies. 15 Q. Yes. Obviously, it's one of many that you 16 cite in your report. 17 Let's in particular shift to the next 18 page where you start really going into the studies that 19 support the link between sex and student evaluations, 20 particularly on Bullet Point 70- -- I'm sorry -- 21 Footnote 75. Are you -- Footnote 75 is an article by a 22 researcher named Boring. 23 A. Uh-huh. 24 Q. Do you know where the -- Gender Biases in 25 Student Evaluations, do you know where the evidence in</p>	<p style="text-align: right;">125</p> <p>1 Q. Okay. And it does discuss that at some point, 2 the specific classes somewhere down here in the 3 analysis. Let's see if we can get there. 22,000 4 observations involving 4,000 students and 372 teachers. 5 It says, "Almost all students are 18 years old, as the 6 first-year undergraduate studies at this university are 7 only open to students who just completed high school." 8 Do you see that? 9 A. Yes. 10 Q. All right. Do you have any understanding that 11 the classes that were reviewed here were STEM classes? 12 A. Again, I'd have to review it to see what 13 the -- no, I can't recall offhand -- 14 Q. Okay. 15 A. -- what the classes were, what they included. 16 Q. And do you have any concerns about relying 17 on studies from Europe about making conclusions 18 about gender bias and student evaluations in the 19 United States? 20 A. To the extent that they tend to be consistent 21 with other studies, I don't have a particular concern 22 here. 23 Q. Okay. Let's look at one of those studies. 24 Let's look at Footnote 76. That's the article by 25 MacNeill. I'll pull that one up.</p>

<p>162</p> <p>1 to kind of give an estimate of that.</p> <p>2 Q. Okay.</p> <p>3 A. But, again, I will just say there's always</p> <p>4 going to be exceptions, absolutely.</p> <p>5 Q. And, similarly, on, like, page 16 you have a</p> <p>6 statement, "Research shows that people tend to excuse</p> <p>7 men's emotional or situational" -- I'm sorry. "Research</p> <p>8 shows that people tend to excuse men's emotionality as</p> <p>9 situational while attributing the same emotional display</p> <p>10 by a woman to an underlying disposition," citing</p> <p>11 Footnote 48, the Barrett case -- the Barrett study. And</p> <p>12 I just want to make sure I'm not missing anything. Do</p> <p>13 you have any understanding of a way to say how much --</p> <p>14 how likely is it that people will tend to do that, based</p> <p>15 on the study?</p> <p>16 A. Again, I could try to figure that out by going</p> <p>17 back to the article, seeing if they report statistics in</p> <p>18 a way that kind of is amenable to that, to giving some</p> <p>19 sense of, you know, how overlapping these distributions</p> <p>20 are between how men and women are rated, to give some</p> <p>21 idea of -- you know, that -- that might be accessible</p> <p>22 to a jury about -- about that effect.</p> <p>23 Q. Okay. And is it generally going to be the</p> <p>24 case that unless you reported a difference -- which, I</p> <p>25 mean, to be fair, on the Joshi study you did report the</p>	<p>164</p> <p>1 Q. (BY MR. GIBSON) If there are others that are</p> <p>2 like: Oh, yeah, I know --</p> <p>3 A. Yeah.</p> <p>4 Q. -- this study or that study, I would be very</p> <p>5 interested; but I'm not expecting you to.</p> <p>6 A. Right. Yeah, I -- not -- as I just sort of</p> <p>7 leaf through them -- oh, on Moss-Racusin, I remember</p> <p>8 that they -- that the effect sizes, I think, on theirs</p> <p>9 were various measures, they said, were moderate to</p> <p>10 large, I -- I believe.</p> <p>11 Q. What --</p> <p>12 A. I believe I might have mentioned that.</p> <p>13 Q. And what footnote are you referring to?</p> <p>14 A. Well, it's 81; but I think that might not be</p> <p>15 the slot. Oh, yeah, yeah, yeah. There it is. "Effect</p> <p>16 sizes for discrimination were moderate to large."</p> <p>17 Moderate to large would mean they were probably around</p> <p>18 a d of .5 to .8, or maybe bigger; and then that would</p> <p>19 be -- I don't know -- you know, maybe from moderate,</p> <p>20 be like maybe two-thirds, to -- to large, maybe like</p> <p>21 80 percent, you know, something like that.</p> <p>22 Q. Okay. All right. I think now I'd like to</p> <p>23 turn to page 40 of your report regarding application to</p> <p>24 the current case and opinions.</p> <p>25 A. We're making progress, aren't we?</p>
<p>163</p> <p>1 14 to -- you know, 14 times more likely. Other than</p> <p>2 that, I don't see a lot of information like that</p> <p>3 included in your report. Is it going to be -- is the</p> <p>4 answer likely to be the same if I ask you about other</p> <p>5 studies, that you would have to look at the study to see</p> <p>6 whether they report effect size to determine whether or</p> <p>7 not you can provide the type of information I'm asking?</p> <p>8 A. Probably. Some of them I might remember,</p> <p>9 like, you know, for instance, I, you know, remembered</p> <p>10 the effect size of the Joshi, et al.; but, yeah, I mean,</p> <p>11 for many of them, probably I would have to go back and</p> <p>12 look.</p> <p>13 Q. Do you remember any of the effect sizes for</p> <p>14 any of the other cases off the top of your head that you</p> <p>15 cited?</p> <p>16 A. Other --</p> <p>17 (Simultaneous speakers.)</p> <p>18 Q. I'm not expecting you to.</p> <p>19 A. Yeah, I don't know.</p> <p>20 Q. I'm just -- you did remember Joshi.</p> <p>21 A. I don't know.</p> <p>22 (Simultaneous speakers.)</p> <p>23 MR. SCHMIDT: I'm going to throw in an</p> <p>24 objection to form. I'm throwing an objection to form</p> <p>25 into that question.</p>	<p>165</p> <p>1 Q. Yes, we are.</p> <p>2 So one of the things that you -- the way</p> <p>3 you describe some of your opinions about how you apply</p> <p>4 the research to this case is that you testify that</p> <p>5 certain behaviors, particularly by Dean Wood, are</p> <p>6 consistent with bias. Did you determine whether they</p> <p>7 are also consistent with no bias? Did you -- did you do</p> <p>8 the analy- -- any analysis to determine whether or not</p> <p>9 her behaviors could be consistent with a lack of bias?</p> <p>10 A. Well, what I'm saying is if you look at some</p> <p>11 of the -- the -- for instance, the content of the</p> <p>12 judgments made of Dr. Nikolova, you know, those are --</p> <p>13 saying that she's not as committed is the kind of thing</p> <p>14 that would occur had motherhood and pregnancy bias</p> <p>15 occurred. So that's what I mean that -- you know,</p> <p>16 consistent with bias.</p> <p>17 What I'm saying is I cannot rule out</p> <p>18 that other explanations, for instance, that, you know,</p> <p>19 Dean Nikolova [sic] you know, simply judged -- she would</p> <p>20 judge a male -- imagine there were a male Dr. Nikolova,</p> <p>21 right? We don't know how she would have judged him. So</p> <p>22 what I'm saying is I can't rule that out that, you know,</p> <p>23 she would have acted the same way against a man.</p> <p>24 But I can say, well, based on the</p> <p>25 research, if she were to discriminate, this is the kind</p>

<p>166</p> <p>1 of form it would take and then look at the case and say:</p> <p>2 Okay. Well, she questioned Dr. Nikolova's commitment,</p> <p>3 which is consistent with the form motherhood and</p> <p>4 pregnancy and taking workplace, you know, accommodation</p> <p>5 policies would -- would tend to take.</p> <p>6 So, again, you know, I'm leaving this</p> <p>7 open. The jury needs to -- to make that ultimate</p> <p>8 decision because I can't say with a scientific</p> <p>9 certainty. I can say: Look, if you look at it, this</p> <p>10 resembles what we'd expect based on the research; but I</p> <p>11 can't tell that, you know, with some sort of scientific</p> <p>12 probability that that was what happened.</p> <p>13 Q. When you say that you expect it based on the</p> <p>14 research, you mean that it -- that it occurs more often</p> <p>15 than not?</p> <p>16 A. I'm not saying -- I'm not making it a more-</p> <p>17 often-than-not judgment because I think that would</p> <p>18 tread on the jury's purview; and I don't want to stick</p> <p>19 a -- you know, a certain probability judgment, which I</p> <p>20 think, you know, that's -- I can't put that kind of</p> <p>21 probability -- I don't feel comfortable putting that</p> <p>22 kind of probability judgment on it.</p> <p>23 If I were a member of the jury, then</p> <p>24 fine, yeah, I've got to make that judgment; and I would</p> <p>25 feel fine making that judgment. You know, I think it's</p>	<p>168</p> <p>1 workplace policies for the purpose of childrearing,</p> <p>2 that those are -- that when you experimentally</p> <p>3 manipulate those things, right -- so you're looking for</p> <p>4 causality -- that what you see is that, on average,</p> <p>5 people will view a woman as less committed to her job</p> <p>6 and that that translates, also, into people viewing</p> <p>7 women as less promotable.</p> <p>8 So these are the biases that the</p> <p>9 research shows occur. And so if you were asking ten of</p> <p>10 my colleagues: Okay. If somebody was pregnant, took</p> <p>11 the -- you know, the leave, all of those things, on</p> <p>12 what dimensions of evaluation would this person</p> <p>13 potentially suffer, suffer discrimination, they'd say,</p> <p>14 "Oh, well, she'd be seen as less committed." Right?</p> <p>15 That would be -- that's, like, a very clear finding from</p> <p>16 the research.</p> <p>17 You know, I can't -- I can't translate</p> <p>18 that into saying that that is the reason why Dean Wood</p> <p>19 made that judgment. What I'm saying is: The research</p> <p>20 suggests that's the kind of judgment that would be made</p> <p>21 if somebody was exhibiting a kind of pregnancy,</p> <p>22 motherhood, workplace accommodation bias.</p> <p>23 Q. But, to be clear, you don't have any idea</p> <p>24 whether or not Dean Wood is exhibiting a pregnancy/</p> <p>25 motherhood bias?</p>
<p>167</p> <p>1 a difficult job the jurors have; but I would -- you</p> <p>2 know, I would be empowered to make that judgment. I</p> <p>3 don't want -- I'm trying not to, yeah, to say something</p> <p>4 that I -- I don't -- that I feel is in any way going to</p> <p>5 mislead the jury.</p> <p>6 Q. Okay. Going to page 41, you state, quote --</p> <p>7 at the bottom of the first full paragraph, "Research</p> <p>8 into biases related to pregnancy, motherhood, and use of</p> <p>9 workplace flexibility policies provide more specific</p> <p>10 information about the pattern of discrimination, if it</p> <p>11 occurred, would likely take." When you say "would</p> <p>12 likely take," how likely is "likely"?</p> <p>13 A. Well, yeah, I guess I could get rid of</p> <p>14 "likely" there. You know, this is a pattern that we</p> <p>15 would expect.</p> <p>16 Q. When you say it's a pattern you would expect,</p> <p>17 is that because there have been studies that confirm</p> <p>18 that that conduct -- or that show that that conduct is</p> <p>19 associated with a statistically significant bias against</p> <p>20 working mothers?</p> <p>21 A. There's research -- if I understand your --</p> <p>22 I'm just going to phrase it a little bit more the way I</p> <p>23 want to phrase it as I'm trying to understand your</p> <p>24 question. So what I would say is there's research that</p> <p>25 shows that pregnancy, motherhood, taking flexible</p>	<p>169</p> <p>1 A. I can't say with any sort of scientific</p> <p>2 certainty. I'm saying that what she said about</p> <p>3 Dr. Nikolova, you know, it's consistent with what we'd</p> <p>4 expect from somebody who is exhibiting this kind of</p> <p>5 bias; but we can't -- I can't rule out that -- that</p> <p>6 those judgments are due to other reasons.</p> <p>7 Q. So how often in the studies that link a</p> <p>8 question -- questioning commitment and competence to --</p> <p>9 or citing to, I guess -- how often in linking those two</p> <p>10 things -- I'm not -- I apologize. I'm not phrasing this</p> <p>11 question very correctly. Let me rethink about how I</p> <p>12 want to phrase the question.</p> <p>13 I'll move on. In the next paragraph</p> <p>14 on page 41, you cite to the fact that Nikolova --</p> <p>15 Dr. Nikolova claims that since she joined the faculty in</p> <p>16 2014, all seven men who went up for tenure and promotion</p> <p>17 received promotion; and three women, including</p> <p>18 Dr. Nikolova, who were considered were denied. Do you</p> <p>19 see that?</p> <p>20 A. Uh-huh. Yes, I see that.</p> <p>21 Q. Do you know if those other individuals were</p> <p>22 mothers?</p> <p>23 A. I do not know if that was the case or not.</p> <p>24 Q. And, in fact, some of the data we've seen has</p> <p>25 shown that if they were not mothers, they actually --</p>

<p>170</p> <p>1 there's data supporting the fact that they may have 2 received a benefit, relative to men, for being nonmother 3 women in the workplace; and they may actually be -- be 4 preferred over men in the workplace. We've seen some 5 data to that effect, correct?</p> <p>6 MR. SCHMIDT: Objection, form.</p> <p>7 A. Oh, you mean the Correll study?</p> <p>8 Q. (BY MR. GIBSON) Yes, that showed that in that 9 actual audit study, nonmother women were significantly 10 more likely to be selected than were the similarly 11 situated men?</p> <p>12 MR. SCHMIDT: Objection, form.</p> <p>13 A. That -- I don't think that -- I don't 14 remember. We'd have to go back whether that comparison 15 was made because I think comparisons were made within 16 gender; but, also, that was a study about callbacks, you 17 know. So it's a little bit different; but, yeah, I 18 mean, I think in this section I'm talking about the -- 19 how mothers, people who take workplace accommodation 20 policies, and pregnant women are treated relative to 21 childless women and to men.</p> <p>22 Q. (BY MR. GIBSON) So you made two points I want 23 to focus on. Number 1, you were, like, the point is 24 childless women compared to women who are mothers and 25 compared to men; but you don't know whether any of the</p>	<p>172</p> <p>1 a hiring decision is a step away from a promotion 2 decision, correct?</p> <p>3 A. That's correct, too. Yes, that's correct, 4 too.</p> <p>5 Q. So you had concerns about citing that study 6 because it was in the context of a callback. Do you 7 have any studies that are in the context of promotion of 8 tenured faculty?</p> <p>9 A. I have studies that were about organizational 10 rewards, such as promotion, like the Joshi 11 meta-analysis. And I cited the Correll study.</p> <p>12 Q. But, again, do you have any studies that are 13 about promotion, tenure and promotion at universities?</p> <p>14 A. Offhand, I'd have to -- I'd have to think 15 about that. Offhand, I can't give you a study about -- 16 that was directly about promotion at universities. 17 There is a -- there is a reference I cite, but I think 18 it's a book about promotion at universities. But, yeah, 19 offhand, I can't think of which ones might be about 20 promotion at a university.</p> <p>21 Q. Kind of at the bottom of page 41 you say, 22 "Should case decision-makers find this information 23 credible, it suggests background conditions that make 24 discrimination more likely."</p> <p>25 A. Sure.</p>
<p>171</p> <p>1 other three women that reportedly did not get tenure had 2 children or not?</p> <p>3 MR. SCHMIDT: Objection, form.</p> <p>4 A. I'm not aware of the parental status of the 5 other women.</p> <p>6 Q. (BY MR. GIBSON) Did you ever ask for that 7 information?</p> <p>8 A. I did not ask for that information, but I'd be 9 happy to review information that -- that could be 10 provided to me.</p> <p>11 Q. And you also pointed out that that study, the 12 Correll study, that actually shows that women without 13 children were preferred over men in getting a callback, 14 was in the con- -- was in the limited context of a 15 callback situation, correct?</p> <p>16 A. Right.</p> <p>17 Q. So would you -- did you make that statement 18 because that means that you're concerned that it doesn't 19 necessarily apply to a tenure promotion decision?</p> <p>20 A. Well, I'm saying that it's -- it's a 21 significant -- there are significant differences with a 22 callback that's not the ultimate hiring decision, right?</p> <p>23 So, you know, that's -- that's a step along the way.</p> <p>24 Q. Do you have any -- well, and you're right. 25 That's not a hiring decision nor is it a -- it's also --</p>	<p>173</p> <p>1 Q. More likely than what?</p> <p>2 A. More likely than if there had been a history 3 of, say, equal promotion of men and women.</p> <p>4 Q. Do you know what the history is of promotion 5 of men and women?</p> <p>6 A. Well, I think there's an allegation here. As 7 I say, Dr. Nikolova alleges that only 4 of 53 tenured 8 faculty within the engineering department are women. So 9 that's a very small proportion of women, and we also 10 know in the research that male-dominated fields tend to 11 be less hospitable, right, to women. And -- and she 12 makes a claim about people going up for promotion. If 13 that -- those claims are incorrect and you want to 14 present evidence that they're incorrect, I'd be happy to 15 review that and this would modify, essentially modify 16 what I say in this section.</p> <p>17 Q. Would you want to look at data that -- well, 18 who's the primary decisionmaker that your report, 19 particularly this section of the report, refers to, 20 Dean Wood, correct?</p> <p>21 A. Yeah, well, Dean Wood is the -- the level of 22 decisionmaking where earlier positive recommendations 23 were overturned, yes. So I think necessarily in this 24 case if there was discrimination, that's kind of the 25 starting point.</p>

<p>178</p> <p>1 the dean's conduct was consistent with discrimination.</p> <p>2 In the Mullenix case, the dean, in fact,</p> <p>3 followed the recommendations of the Budget Committee and</p> <p>4 you still criticized the dean for following the</p> <p>5 recommendations of the Budget committee --</p> <p>6 MR. SCHMIDT: Objection, form.</p> <p>7 Q. -- and you did not point out that, in that</p> <p>8 opinion, that the Budget Committee itself was a form of</p> <p>9 mitigating bias by the fact that it was a committee of</p> <p>10 multiple people, as you do in this case. Is there a</p> <p>11 reason why you didn't do that in the Mullenix case but,</p> <p>12 yet, you do it here?</p> <p>13 MR. SCHMIDT: Objection, form.</p> <p>14 A. I think my -- my opinions in both cases are</p> <p>15 entirely consistent if you carefully read my report. In</p> <p>16 my report in the current case for Dr. Nikolova, what I</p> <p>17 talk about is just what we were talking about. The</p> <p>18 Budget Council in this case is in a different</p> <p>19 department, a different school than the Budget Council</p> <p>20 in the other case.</p> <p>21 The Budget Council in this case had</p> <p>22 carefully-constructed benchmarks. For instance, they</p> <p>23 had looked at recently-tenured individuals at peer</p> <p>24 institutions and looked at things like the citation</p> <p>25 counts for their research, their grant dollars that they</p>	<p>180</p> <p>1 institutions. They should come up with the same</p> <p>2 recommendation.</p> <p>3 By contrast, the other Budget Council was</p> <p>4 much more along the lines of very subjective potential</p> <p>5 decisionmaking where it was kind of one person's opinion</p> <p>6 one year, another person's opinion the next year, right,</p> <p>7 those sorts of things that could be influenced by common</p> <p>8 biases.</p> <p>9 Let's say that I'm a Budget Council</p> <p>10 member in the Nikolova case and let's say, you know, I</p> <p>11 have some bias against her; but if I'm charged with I</p> <p>12 have to look up her citation count and compare it to</p> <p>13 these already-set benchmarks, well, then, you know, the</p> <p>14 conclusion is pretty much going to be based on, you</p> <p>15 know, what the -- what I'm instructed to do. I don't</p> <p>16 have the room to -- to allow -- you know, I don't have</p> <p>17 the room to introduce my bias in that way.</p> <p>18 MR. GIBSON: I'd like to take just a</p> <p>19 quick five-minute break.</p> <p>20 THE WITNESS: Okay.</p> <p>21 THE VIDEOGRAPHER: The time is 3:06 p.m.</p> <p>22 We are off the record.</p> <p>23 (Off the record from 3:06 to 3:17 p.m.)</p> <p>24 THE VIDEOGRAPHER: The time is 3:17 p.m.</p> <p>25 We are back on the record.</p>
<p>179</p> <p>1 had obtained, the impact factor of the journals that</p> <p>2 they published in. All of these sorts of things were --</p> <p>3 were very carefully laid out, very carefully assessed.</p> <p>4 Now, again, some of those metrics --</p> <p>5 there's some bias creeping into some of those metrics;</p> <p>6 but once you look at the Budget Council level, the</p> <p>7 Budget Council in the engineering school is completely</p> <p>8 different in their procedures than the other Budget</p> <p>9 Council you're talking about, who would, for instance,</p> <p>10 assign one person to read the target individual's work</p> <p>11 and render an opinion, right, which leaves open the door</p> <p>12 to all sorts of subjectivity.</p> <p>13 On this Budget Council -- let's put it</p> <p>14 this way: You could take the engineering -- the Budget</p> <p>15 Council that evaluated Dr. Nikolova, if -- if you set</p> <p>16 out -- set out all their criteria, right, all of the</p> <p>17 things that they used, how they weighed them, the</p> <p>18 benchmarks they compared them to, then you could just</p> <p>19 bring in a whole new set of individuals, a whole new set</p> <p>20 of faculty who adhere to the same standards, they're</p> <p>21 going to find that Dr. Nikolova, when they research her</p> <p>22 citation count, they're going to get the same number,</p> <p>23 right? That's not going to be a subjective judgment on</p> <p>24 their part. They're going to compare that to the same</p> <p>25 benchmark of recently-tenured individuals at other</p>	<p>181</p> <p>1 MR. GIBSON: I think Robert would like to</p> <p>2 make a --</p> <p>3 (Simultaneous speakers.)</p> <p>4 MR. SCHMIDT: I'm sorry. Go ahead,</p> <p>5 Darren.</p> <p>6 MR. GIBSON: No.</p> <p>7 MR. SCHMIDT: Since we're back on the</p> <p>8 record, I would request that Dr. Glick be given a copy</p> <p>9 of his deposition to read and sign and make corrections</p> <p>10 if necessary. That is what I have.</p> <p>11 MR. GIBSON: Okay.</p> <p>12 Q (BY MR. GIBSON) Dr. Glick, on page 42 of your</p> <p>13 report, you referred -- we were mentioning the fact that</p> <p>14 there's a reference to comparator information regarding</p> <p>15 other male faculty who were promoted. Have you done any</p> <p>16 separate analysis of the information regarding</p> <p>17 comparators in this case?</p> <p>18 A. No, I did not.</p> <p>19 Q. Excuse me. I'd like to go to your statements,</p> <p>20 I think -- oh, let's first go to page 44, please.</p> <p>21 First, you say -- you're talking about evaluations. You</p> <p>22 say, "Thus, being visibly pregnant may well have</p> <p>23 unfairly affected Dr. Nikolova's teaching evaluations."</p> <p>24 Do you know if Dr. Nikolova was visibly pregnant at the</p> <p>25 time of the evaluations in question?</p>

<p>198</p> <p>1 A. Yes, I cannot -- this is like saying can 2 you -- you know, you might be able as a meteorologist to 3 predict the pattern of weather with a certain degree of 4 accuracy; but are you going to be able to predict the 5 path of a blizzard in a snowflake? No. When you're 6 asking can I predict what an individual is going to do, 7 you know, in these kinds of circumstances, that's -- you 8 know, that's -- that would be very, very dicey. 9 Q. It's not only you can't predict you it; you 10 can't, after the fact, determine whether or not specific 11 decisions were motivated -- motivated by specific 12 factors. You can't make a scientific determination as 13 to whether or not Dr. Wood -- Dr. Wood's gender and 14 prior experience, in fact, made her more likely to 15 discriminate or, in fact, caused her to discriminate? 16 A. As an individual, right. I mean, we already 17 went over this, I think. I think I already answered 18 that question. And if I understand the question 19 correctly, the answer is: No, I cannot. And I've 20 already been very upfront about that and -- and, you 21 know, that my whole section about application to the 22 case is constantly framed with that kind of caveat. 23 Q. Let's see. It'll take me just a couple of 24 minutes here. Hold on just a second. I've got to 25 transition.</p>	<p>200</p> <p>1 with potential bias and discrimination? 2 A. So, again, I think you're asking me if the 3 final section of my report involved doing a formal 4 scientific study on the University of Texas and the 5 people involved. And the answer is: No, I did not. 6 Q. Are you aware of any scientific technique that 7 would allow -- that would support statements, the 8 statements that are found in the -- your section 9 regarding the application of this case as to whether 10 Dean Wood's conduct was consistent with discrimination? 11 A. So, again, I'm applying my expertise in the 12 scientific framework to offer avenues for the jury to 13 consider and to make observations about the case; but 14 that section of the report, as I made clear in the 15 report, is not, itself, performing a scientific study. 16 And I, therefore, resist making precise scientific 17 conclusions as a result. 18 Q. And that methodology that you -- or the 19 opinions that you provide in Section V regarding 20 application to this case are not based on methodologies 21 that have been subject to peer review or publication; is 22 that correct? 23 A. That -- yes, that would be correct. I think 24 that's basically what I was trying to say. And I think 25 your other question was, you know, is there a method,</p>
<p>199</p> <p>1 Okay. I apologize. Hold on just a 2 second while I pull something up. 3 MR. SCHMIDT: While you're doing that, 4 Debbie, do you have a time in terms of how long we've 5 been going? 6 THE REPORTER: Four hours, twenty-four 7 minutes. 8 MR. SCHMIDT: Thanks. 9 I know Darren's going to be done by 4:00, 10 so. 11 MR. GIBSON: I just -- I think I'm going 12 to wrap up. I think these are pretty easy. 13 MR. SCHMIDT: I'm joking. 14 MR. GIBSON: Oh. 15 MR. SCHMIDT: But go ahead. I was giving 16 you a hard time. 17 MR. GIBSON: I'm sorry. I wasn't 18 listening. I was focused on something else, but I think 19 I'm going to be pretty close. 20 Q (BY MR. GIBSON) I'm going to ask you a few 21 questions. I just want to make sure that -- I think 22 these are relatively easy questions. 23 Is there a scientific technique that 24 supports the rendering of your judgment in this case as 25 to whether or not what occurred here would be consistent</p>	<p>201</p> <p>1 you know, to determine, for instance, with a scientific 2 certainty, whether discrimination occurred in a 3 particular case. There isn't a method to do that, at 4 least not for an individual case such as this. 5 Q. Okay. Do you have a personal opinion as to 6 whether or not Dr. Nikolova was subject to 7 discrimination in the tenure decision that's at issue in 8 this case? 9 A. I don't want to render any personal opinion on 10 this because I think that's going outside my role. 11 Q. That's fine, but do you have a personal 12 opinion? 13 A. Having reviewed all of the materials in the 14 case, if I were a juror, sure, I have an opinion on it. 15 Q. Based on looking at the evidence that you've 16 reviewed, what's your personal opinion? 17 MR. SCHMIDT: Objection, form. 18 A. I'm not going to state my personal opinion 19 because I think that's not my role. I think that would 20 be -- that would be -- to me, that's painting outside 21 the lines of what I understand my role to be. 22 Q. (BY MR. GIBSON) Do you think your personal 23 opinion may have affected the conclusions that you've 24 reached in your report? 25 MR. SCHMIDT: Objection, form.</p>

Journal of Management
Vol. XX No. X, Month XXXX xx-xx
DOI: 10.1177/0149206310374774
© The Author(s) 2010
Reprints and permission: <http://www.sagepub.com/journalsPermissions.nav>

A Meta-Analysis of Gender Group Differences for Measures of Job Performance in Field Studies

Philip L. Roth

Clemson University

Kristen L. Purvis

Cornell University

Philip Bobko

Gettysburg College

There are multiple views in human resource management and organizational behavior concerning gender differences in measures of job performance. Some researchers suggest that males generally are evaluated higher than females across a variety of situations that include job performance measurement. At the same time, the presence of specific status cues in expectation states theory (EST; similar to the concept of individuating information) suggests that measures of job performance will be more similar than different for males and females. Previous analyses are unclear in their results for the measurement of the construct of job performance because they have included, and/or focused on, additional constructs (e.g., hiring suitability, leadership performance aggregated with leadership satisfaction) or have used student samples in lab experiments. The authors of this article conducted a meta-analysis of job performance measures from field studies. They found that females generally scored slightly higher than males (mean $d = -.11$, 80% credibility interval of $-.33$ to $.12$). Other analyses suggested that, although job performance ratings favored females, ratings of promotion potential were higher for males. Thus, ratings of promotability may deserve further attention as a potential source of differential promotion rates. These findings and processes are discussed within the context of EST.

Keywords: gender; job performance; meta-analysis

Corresponding author: Philip L. Roth, Department of Management, Clemson University, Clemson, SC 29634-1305, USA

Email: rothp@clemson.edu

Introduction

Job performance is an important variable in the human resource management and organizational behavior literatures (e.g., Viswesvaran, Schmidt, & Ones, 2005). One important issue in measuring job performance is whether or not there are stable, persistent differences between ethnic and gender subgroups. While there are two recent meta-analyses relating ethnicity to job performance (McKay & McDaniel, 2006; Roth, Huffcutt, & Bobko, 2003), there has been no recent meta-analytic attention to gender differences in job performance (McKay, 2009). This is unfortunate, as there are multiple views on this issue. One such view is noted in an often-cited narrative review article by Nieva and Gutek (1980: 273). These authors suggested that there is a “fairly consistent” tendency for males to receive more favorable performance evaluations than females receive (see also Greenhaus & Parasuraman, 1993; Igbaria & Baroudi, 1995; Lyness & Heilman, 2006). In turn, higher evaluations of job performance for males could have serious consequences in promotion decisions (Greenhaus & Parasuraman, 1993; Mobley, 1982) and might be a precursor of the glass-ceiling effect (Heilman, 2001; L. M. Shore & Thornton, 1986; T. H. Shore, 1992).

At the same time, there can be a great deal of performance-related information (i.e., performance cues) available to individuals who assess job performance. Expectation states theory (EST) suggests that these specific cues are likely to have a much greater influence on measures of job performance (e.g., ratings) than do cues such as gender. Similarly, there are many psychological similarities between genders, as well as data supporting that view (e.g., Hyde, 2005). Thus, an alternative view is that assessments of male and female job performance may be more similar than different.

Previous meta-analyses have begun to address gender differences on variables related to job performance, but not the measurement of job performance per se. Some meta-analyses have focused on experimental studies of hiring recommendations (Davison & Burke, 2000; Olian, Schwab, & Haberfield, 1988). Others have cumulated mixes of laboratory and field studies while aggregating heterogeneous outcome variables such as leadership effectiveness, satisfaction with leaders, and perceived leadership ability (Eagly, Karau, & Makhijani, 1995; Eagly, Makhijani, & Klonsky, 1992). The different methods, samples (e.g., students vs. incumbents), dependent variables, and mixed results make it difficult to determine if, overall, males and females differ on measures of job performance in operational field settings.

The purpose of this article is to meta-analyze male–female differences for measures of job performance in field settings (e.g., supervisory ratings or measures of output).¹ We do not address differences in salaries, an outcome of performance that has already been meta-analyzed (Ng, Eby, Sorensen, & Feldman, 2005). Instead, we focus on determining if available evidence supports gender similarity in actual or rated levels of job performance in field settings.

We examine “direct” measures of job performance (Cascio, 1997). Such measures include supervisory ratings, measures of quantity and quality, and measures of output, rather than “indirect measures” such as withdrawal behaviors (e.g., absenteeism). We also added measures of promotion potential, because several primary studies contrasted this measure to measures of job performance.

The Importance of Gender Differences in Job Performance

Differences between males and females, on a variety of variables including the measurement of job performance, have received much attention in both the popular press (e.g., MacGillivray, Beecher, & Golden, 2009) and the scientific literature (e.g., Hyde, 2005; King, Hebl, George, & Matusik, 2010). The topic is important for several reasons. In addition to the reasons already noted, the actual relationship between gender and measures of job performance is important for evaluating theories. For example, a widely cited narrative review suggests widespread differences in ratings across gender, such that females are generally disadvantaged (Nieva & Gutek, 1980). Others have suggested that gender incongruity with the perceived role of managers could lead to lower ratings for female managers (e.g., Eagly & Karau, 2002; Schein, 2001) or that females are not perceived as positively as males are in terms of promotability (Lyness & Heilman, 2006).

Gender differences in measures of job performance can influence many other variables within organizational settings. Measured job performance differences could influence both pay and promotions (Robertson, 1986). Or, unjustified differences in performance measures could erode trust (Schoorman, Mayer, & Davis, 2007) or lead to lower levels of job satisfaction (as per Colquitt, Conlon, Wesson, Porter, & Ng, 2001; see also King et al., 2010), which, in turn, lead to withdrawal such as lateness, absenteeism, and turnover (Harrison, Newman, & Roth, 2006). Accurate assessment of the magnitude of any gender differences in the measurement of job performance is also important for the application of human resource models of selection test fairness (Cleary, 1968, see also Bartlett, Bobko, Mosier, & Hannan, 1978). Further, managers deal with a wide variety of demands each day (e.g., Mintzberg, 2001) and may benefit from guidance on where to focus their efforts at creating a fair and balanced workplace.

Previous Meta-Analyses From Related Literatures

Several different disciplines contain literature that contributes to our understanding of, and expectations for, gender group differences in measures of job performance. We briefly review key meta-analyses from the various disciplines.

Experimental studies in personnel selection. Some studies (primarily in applied psychology and judgment and decision making) cumulate experimental studies of hiring decisions. Olian et al.'s (1988) meta-analysis focused on laboratory studies of individuals making decisions or evaluations of hypothetical candidates (see also Davison & Burke, 2000). Olian et al. reported a mean d of .41 ($K = 19$, $N = 1,842$), such that males were evaluated, on average, higher than females for hiring-related ratings. Olian et al. also examined the effect of qualifications on hiring-related judgments and found that hiring qualifications accounted for approximately eight times as much variance in the evaluations as did their gender manipulations (more on this below, as we link this analysis to the concept of "individuating information").

Leadership. Two meta-analyses in leadership focused on gender differences. The first meta-analysis focused primarily on laboratory studies (mainly using written or video vignettes

of leaders; see Eagly et al., 1992, including their analysis of the article by Swim, Borgida, Maruyama, & Myers, 1989). The second leadership meta-analysis cumulated lab and field studies (Eagly et al., 1995). The primary studies in this analysis used objective measures of leader performance, including performance in laboratory studies, measures of job knowledge (e.g., tacit knowledge of business), and absenteeism. Subjective measures included self, subordinate, and peer ratings to assess leader effectiveness, leader motivation, leader ability, and satisfaction with the leader (Eagly et al., 1995), although self-report measures by leaders and measures using subordinate reports were most prevalent (see their Table 3). Overall results showed a d of $-.02$ ($K = 76$).

Applied psychology and education. Bowen, Swim, and Jacobs (2000) cumulated measures of male–female differences using studies in applied psychology and education that reported information on covariates of the gender–performance relationship (e.g., organizational level, experience, education). The overall mean d was $-.01$ ($K = 32$, N not reported). Another set of analyses isolated studies in which there were no self-report measures (mean $d = -.05$, $K = 22$, N not reported). The authors interpreted their results as suggesting little overall bias for gender in performance appraisals.

In applied psychology, a large-scale empirical study (Sackett, DuBois, & Noe, 1991) focused on personnel selection within a single, multiple-job, nonstudent database used to evaluate the validity of the General Aptitude Test Battery (GATB). The researchers created 486 firm–job combinations in which they examined male–female differences based on supervisory performance appraisals, and they found that males were rated higher than females ($d = .07$).

Limitations of previous literatures. Overall, there are a variety of reasons why it is difficult to ascertain an accurate level of gender differences in measures of job performance in organizational settings. First, many results may be limited by a lack of individuating information and a lack of generalizability to actual measures of job performance in organizational settings (Davison & Burke, 2000). This concern applies most clearly to the experimental studies using student participants in personnel selection but is also relevant to the leadership studies (as student samples in laboratory studies were present in both meta-analyses).

Second, some meta-analyses, particularly in the leadership literature, contain a highly heterogeneous set of dependent constructs or variables that includes ratings of leader performance, leader satisfaction, job knowledge, absenteeism, and so on. Thus, it is difficult to isolate the gender–job performance relationship. Relatedly, many jobs may not involve leadership-related tasks and duties, so generalization of this literature to those other jobs is problematic.

Third, several sources of information in the prior analyses are of concern. For example, some ratings in the above analyses are self-ratings, in spite of generalized concern about such ratings (Farh & Dobbins, 1989; Harrison & Shaffer, 1994). Again, this concern is most focused on leadership studies. (In this regard, Bowen et al., 2000, note that more than 90% of performance appraisals in organizations are conducted by supervisors.) Further, in the education literature, we are concerned by use of student ratings of teachers, because such ratings are not generally accepted measures of job performance in human resource management and organizational behavior.

Finally, the fact that managerial and professional jobs were “virtually absent from the database” (Sackett et al., 1991, p. 264) from the GATB study raises concerns. Results based upon the inclusion of many “blue-collar” jobs (e.g., machinist) could readily be confounded by association of the jobs within that database with male stereotypes (see Davison & Burke, 2000).

Theoretical Perspectives and Research Hypotheses

Expectation states theory. EST provides a theoretical framework to understand gender differences in the measurement of job performance (see Berger, Rosenholtz, & Zelditch, 1980; Berger, Wagner, & Zelditch, 1985; Correll & Ridgeway, 2003; Dovidio, Ellyson, & Keating, 1988). EST suggests that gender is one of several “diffuse” (or indirect) status cues that influence expectations about the knowledge, ability, or influence of a given person. Diffuse cues (e.g., gender or ethnicity) can be conceptually linked to performance on many tasks, although the links may not be as strong as with specific status cues. The diffuse cue of gender is generally thought to suggest that females are more artistic and literary, kinder, and more patient and understanding, while men are thought to be more scientific, mechanical, and assertive (e.g., Berger et al., 1980; see also Eagly & Karau, 2002). This can translate into females initially having less status in many work situations (Berger et al., 1985). For example, all else equal, a female in an auto parts store may be judged to have a lower level of job knowledge than a male does.

Specific status cues are typically more directly related to performance expectations than diffuse status cues are. That is, specific cues typically have a fairly short logical chain relating them to expectations (e.g., the cue of spatial ability is directly linked to job performance in air traffic control), whereas diffuse expectations have longer logical chains (e.g., the cue of male might relate to spatial ability, which relates to job performance in air traffic control). When both direct and diffuse status cues are present, EST researchers note that the task-relevant, specific cues usually substantially outweigh the diffuse cues (Dovidio et al., 1988; Freese & Cohen, 1973; W. Wood & Karten, 1986). To continue the above example, the female in the auto parts store may be judged to have superior job knowledge when one learns she has a master’s degree in mechanical engineering and is a certified mechanic.

Similarly, social psychologists and sociologists have used the construct of individuating information to understand judgments about gender differences in the workplace (e.g., Reskin, 2000; see also Eagly et al., 1992, 1995, or Olian et al., 1988). Observed performance on the job in question (and/or job-related knowledge, skills, and abilities) could be a particularly powerful source of individuating information for someone assessing job performance. Reskin notes that individuating information can counter stereotypes, given the relevance of such information to many judgmental tasks (e.g., hiring decisions). This use of individuating information is quite similar to the role of specific cues in EST in that individuating information should be more closely related to performance than are diffuse cues. That is, individuating information may minimize the role of dynamics such as stereotypes.

The application of EST, and the related concept of individuating information, has several implications for *field* studies of job performance. In particular, the availability of substantial

amounts of actual performance information can be seen as a potentially large volume of important specific status cues. Thus, in field studies of job performance, one would posit smaller gender differences in measures of job performance (and differences might slightly favor females, given the material noted below). In contrast, in experimental studies on performance evaluation or hiring potential, specific cues (and their variability) are often controlled by keeping them constant. Thus, EST helps predict a difference in results between field studies of job performance assessment and experimental studies.

In sum, given the above logic using EST and the convergence with the concept of individuating information, we hypothesize:

Hypothesis 1: Gender differences on measures of job performance in field studies will tend to be small (e.g., $d = .2$ or less).

Theories of job performance. The theory of job performance by Campbell and colleagues (e.g., Campbell, Gasser, & Oswald, 1996; Campbell, McCloy, Oppler, & Sager, 1993; McCloy, Campbell, & Cudeck, 1994) suggests that three constructs are the immediate precursors of underlying or actual job performance—declarative knowledge, procedural knowledge, and motivation (see also the work of Schmidt and Hunter, such as Hunter, 1983a; Schmidt, Hunter, & Outerbridge, 1986; Schmidt, Hunter, Outerbridge, & Goff, 1988).

Given the empirical evidence about these three precursors of job performance, there appears to be little that would suggest higher job performance for males across a variety of jobs. For example, regarding declarative job knowledge, a study of post-training measures of job knowledge reported d s of approximately $-.50$ to $-.60$ across several clerical jobs in a military organization (Dunbar & Novick, 1988). Thus, females scored higher than males. Given that declarative or procedural knowledge is a precursor to job performance, this suggests that females might outperform males (although this is only one study).

There also appears to be little reason to posit gender differences in underlying or measured job performance due to motivation. In terms of trait motivation, differences between males and females on personality factors are relatively small. For example, Hough, Oswald, and Ployhart (2001) report a d of $.06$ for achievement motivation (see also Hyde, 2005). Similarly, there are few reported sizeable differences on other personality factors that might influence job performance to the advantage of males (e.g., Hough et al. report d s of $.09$ for extraversion, $-.08$ for overall conscientiousness, $-.22$ for dependability, $.24$ for adjustment, and $-.39$ for agreeableness; N s were more than 20,000 in all cases). Similarly, overall cognitive ability differences were generally minimal for males versus females (Hough et al., 2001).

Thus, females tend to score higher on the job knowledge variables and fairly similarly, on average, on other variables, such as trait motivation, personality, and cognitive ability. So, available data on precursors of job performance suggest that females might perform better than males on the job. Conversely, it also appears there is no reason to generally attribute higher underlying performance to males across a variety of jobs.

Promotability. In contrast to job performance ratings, promotability ratings are expected to favor males for a number of reasons. For example, females can receive less credit for their contributions (Hielman & Haynes, 2005). Likewise, female managers or leaders may have

to deal with more negative attitudes and may find that they have less access to leadership positions (e.g., see the review of role congruity theory by Eagly & Karau, 2002; see also Eagly et al., 1995).

EST suggests that promotion decisions may be more open to the influence of diffuse status cues because the process of considering the *predicted* level of performance in a higher level job involves more judgmental uncertainty than does the process of rating the *actual* performance of an incumbent in a given job. So, promotion decisions, by their nature, often do not allow decision makers maximal access to performance information (because most candidates have not done the job in question), thereby creating more opportunities for increased weight to diffuse cues.

Thus, we hypothesize:

Hypothesis 2: Males will score higher than females on promotability ratings.

Moderators. We also examined several moderators with exploratory analyses. For example, we coded time (i.e., date of study) to see if gender differences were relatively stable across decades due to factors such as the influence of specific cues (as per EST) or if *ds* changed across decades, as might be suggested by the potentially weakening of stereotypes or other related mechanisms (Nieva & Gutek, 1980). We also examined other potential moderators such as the purpose of ratings (as per Landy & Farr, 1980) to see if ratings for administrative purposes were associated with *d* values less than ratings made for research purposes (e.g., Tsui & O'Reilly, 1989). We also conducted similar analyses for job complexity and publication status (published vs. nonpublished documents).

Method

Literature Search

We searched databases for articles that reported data on measures of job performance effect size and gender. We searched PsycINFO (from the American Psychological Association), Business Source Premier, General Business File, Educational Resources Information Center (also known as ERIC), and Dissertation Abstracts International. We also checked the reference lists of previous analyses that focused on field data (e.g., Bowen et al., 2000). Finally, we wrote to researchers active in the area both for the availability of additional studies and to clear up any questions or get more detailed information.

Inclusion Criteria

There were four criteria necessary for studies to be included in our meta-analysis. First, studies must have reported measures of job performance (e.g., ratings, measures of output). As noted earlier, we did not include studies of salary (e.g., Ng et al., 2005) because we viewed salary as an outcome related to the assessment of job performance.

Second, ratings must have been completed by supervisors or peers. Thus, we did not include studies using self-report measures of job performance (e.g., Jagacinski, 1987), nor did we include student reports of teacher performance (as done by Bowen et al., 2000).

Third, participants being assessed must have been job incumbents. Studies asking students to rate paper or videotaped people were not included (as per studies in Davison & Burke, 2000; Olian et al., 1988).

Fourth, we paid careful attention to data dependence. We computed unit-weighted composites of performance when multiple correlations were present (e.g., Ployhart, Wiechman, Schmitt, Sacco, & Rogg, 2003). We also monitored data in order to avoid double counting similar or overlapping data sets published in different journals (see J. Wood's [2008] discussion of this issue). For example, we coded only one effect size for each of the following pairs of articles (e.g., Greenhaus & Parasuraman, 1993; Greenhaus, Parasuraman, & Wormley 1990; Pulakos & Wexley, 1983; Wexley & Pulakos, 1982). However, we did allow articles with multiple independent data sets to contribute multiple coefficients to our analyses (e.g., Arvey, Miller, Gould, & Burch, 1987; Saad & Sackett, 2002).

Coding

The first two authors, one male and one female, individually coded the data from each study. Measures of interrater agreement on the initial codes (correlation coefficients and percentage agreement) are reported below. Following the individual coding, the two coders met to resolve any discrepancies via consensus.

A number of variables were coded in addition to the sample size and the standardized gender group difference (respectively, interrater $r = .99$ and $r = 1.00$).² Coders first coded the nature of the measure (subjective vs. objective; 100%) and purpose of the measure (administrative, research, counseling, not sure; 96%). Second, the area of job performance was noted (e.g., rating of performance, output, quality or complaints, promotability; 100%). Job complexity was also coded using Hunter's (1983b) codes: low (e.g., mail sorter), low-medium (e.g., truck driver), medium (e.g., skilled crafts), medium-high (e.g., computer troubleshooter), or high (e.g., executives, scientists; $r = .92$). Finally, the date and publication status of the source was recorded (e.g., refereed journal, thesis or dissertation, etc.; 100%). We searched for the dates that data were gathered but found little such information, so we used date of publication for this variable when actual data collection dates were not available.

Analyses

We used the Hunter-Schmidt approach to meta-analysis (Hunder & Schmidt, 2004). Calculations were conducted with the Schmidt-Le program (version 1.1; Schmidt & Le, 2004). We corrected the effect sizes for unreliability in the measure of job performance. This generally involved the use of internal consistency measures of reliability from the performance data in our studies because there were only two studies that reported interrater reliabilities of ratings. Hunter and Schmidt state that using internal consistency measures of

Table 1
**Meta-Analysis of Gender Differences for Measures of Overall
Job Performance and Promotability**

Analysis	Mean <i>d</i>	Corrected <i>d</i>	<i>K</i>	<i>N</i>	Var.	80% CRI	% Art.
Overall job performance	-.10	-.11	61	45,733	.030	-.33 to .12	17
Objective vs. subjective measures							
Supervisory ratings	-.13	-.14	50	36,896	.014	-.29 to .01	31
Objective measures	-.02	-.02	4	4,744	.015	-.18 to .14	20
Objective: no large sample	-.11	-.12	3	1,426	.040	-.37 to .14	20
Promotability ratings							
Promotability	.10	.11	8	4,550	.038	-.14 to .36	18
Matching studies	.11	.12	6	4,307	.035	-.12 to .36	15
promotability							
Matching studies	-.09	-.09	6	4,309	.001	-.12 to -.07	93
performance							
Manager studies							
Managers	-.14	-.15	8	7,555	.004	-.23 to -.07	55
Managers: no large samples	-.08	-.09	6	1,961	.005	-.18 to .002	74

Note: Mean *d* is the observed *d* across studies, corrected *d* is corrected for measurement reliability in job performance, *K* is the number of coefficients, *N* is the number of participants, Var. is the variance of the estimate of the corrected *d*, CRI is an 80% credibility interval, and % Art. is the percentage of variance due to the artifacts of sampling error and measurement unreliability. We also removed police and military studies from these analyses, given their somewhat different pattern of results from other studies (details available from the first author).

reliability is generally expected to produce a conservative correction of effect sizes. For some of our ancillary analyses, we used the reliability value of .6 for measures of promotability. We adopted this value from previous literature (e.g., McKay & McDaniel, 2006; Roth et al., 2003). This value is also similar to the reliability of supervisor ratings of leadership ability ($r = .53$, $K = 10$, $N = 2,171$) in the meta-analysis by Viswesvaran, Ones, and Schmidt (1996).

Results

Overall Results

The results for male versus female differences for measures of job performance are reported in Table 1. Our overall analyses include studies using performance ratings and measures of output (e.g., sales volume), although only three studies in this analysis contained objective measures of performance.³ We do *not* include our measures of promotion potential in these overall analyses, as they are noted separately in the table.

The overall analysis of job performance measures resulted in a mean corrected *d* of -.11 ($K = 61$, $N = 45,733$). This suggests that females, on average, were rated as performing somewhat better than males in operational field settings. The 80% credibility interval ranged from -.33 to .12. This suggests that there is substantial variation in estimated population values and that in some cases males perform better (positive values) and that in a somewhat

greater proportion of cases females perform better (negative values). In general, the data show substantial gender similarity in mean levels of overall job performance.

Overall, the data support our hypothesis (Hypothesis 1) of relatively small differences in measures of operational job performance (and if there are differences, as indicated by the range in credibility values, females generally have slightly higher average ratings or output). The exclusion of large samples does not substantially change the mean estimates.⁴

We also report results of supervisory or subjective ratings versus objective measures (again, see Table 1). Results for supervisory ratings indicate that $d = -.14$ ($K = 50$, $N = 36,896$), with a credibility interval of $-.29$ to $.01$. Objective measures are associated with $d = -.02$ ($K = 4$, $N = 4,744$). Results change somewhat when a large sample is deleted, to $d = -.12$ ($K = 3$, $N = 1,426$). We interpret the latter values with caution because the number of studies and sample sizes are small. Results for objective measures paint a reasonably consistent picture with performance ratings in which gender differences were small (although females score slightly higher on average). Thus, there is some convergence across different types of measures.

Promotability. Several studies reported measures of promotability (e.g., Harris, Kacmar, & Carlson, 2006). As noted in Table 1, the corrected mean d was $.11$ ($K = 8$, $N = 4,550$). Thus, on average, males scored somewhat higher than females on measures of promotability, although the credibility interval was again wide. Interestingly, a subset of six studies also reported measures of performance on the current job (N s differ slightly on these analyses due to missing data in one primary study). Results show a mean d of $-.09$ for job performance ($K = 6$, $N = 4,309$) such that female performance was slightly higher, while at the same time the mean d for ratings of promotability was $.12$ ($K = 6$, $N = 4,307$). Below, we return to this unique finding and analysis (directly comparing the promotability d to the performance d).

We also report results for job performance for manager samples separately, given past attention to this type of job (Eagly et al., 1992; see also Eagly & Karau, 2002, or Schein, 2001). In the bottom panel of Table 1, we also report a mean corrected d of $-.15$ ($K = 8$, $N = 7,555$) for all managerial samples, and we note that the credibility interval does not include zero. The mean d is $-.09$ ($K = 6$, $N = 1,961$) after removing one large sample, and the credibility interval does include zero. Thus, point estimates of gender group differences (i.e., mean d s) suggest more similarity than difference, although sample sizes are small. Additionally, consistent with our overall analyses, the tendency is that for nonzero values the female average is greater than the male average. We also note that these jobs were more representative of middle management positions than of upper management positions.

Moderators

Before examining some of our other moderators, we examined the type of organization associated with the data (see Table 2), based on strong statements in previous meta-analyses. Specifically, Eagly et al. (1995) suggested that “military organizations deviated strongly from all other classes of studies” (p. 135; see p. 138 for further cautions).

Table 2
Meta-Analysis of Time, Publication Status, Job Complexity,
and Purpose of Data Gathering

Analysis	Mean <i>d</i>	Corrected <i>d</i>	<i>K</i>	<i>N</i>	Var.	80% CRI	% Art.
Type of organization							
Industrial	-.12	-.12	41	36,326	.013	-.27 to .02	29
Public sector and educational	-.31	-.33	8	1,147	.045	-.61 to -.07	42
Public sector: no large sample	-.19	-.20	7	769	.019	-.38 to -.02	68
Military/police	.00	.00	12	8,260	.086	-.37 to .37	7
Military/police: no large sample	.06	.07	11	5,276	.123	-.38 to .51	7
Date of publication							
1969-1988	-.14	-.15	18	24,231	.007	-.25 to -.04	34
1969-1988: no large samples	-.18	-.19	16	7,675	.011	-.33 to -.06	46
1989-2001	-.11	-.12	14	3,315	.031	-.34 to .10	39
1989-2001: no large samples	-.14	-.15	13	2,521	.039	-.45 to -.01	49
2002-2009	-.05	-.05	29	18,187	.055	-.36 to .25	11
2002-2009: no large samples	-.03	-.03	27	12,259	.075	-.38 to .32	12
Publication status							
Journal	-.10	-.11	44	26,843	.041	-.37 to .15	15
Not published	-.11	-.12	17	18,872	.012	-.26 to .02	26
Not published: no large samples	-.08	-.09	15	2,316	.071	-.43 to .25	30
Job complexity							
Low and low-medium complexity	-.07	-.07	29	19,404	.049	-.36 to .22	12
Medium complexity	-.12	-.13	14	15,645	.014	-.28 to .02	23
Medium-high complexity	-.13	-.13	9	8,841	.003	-.20 to -.07	64
Medium-high complexity: no large samples	-.06	-.07	7	3,247	.000	-.07 to -.07	100
Purpose of performance measure							
Administrative	-.06	-.07	12	7,341	.024	-.27 to .13	23
Administrative: no large samples	-.13	-.14	11	4,023	.034	-.38 to .10	27
Research	-.10	-.11	43	35,411	.031	-.33 to .12	15
Research: no large samples	-.07	-.08	39	13,434	.077	-.43 to .28	15

Note: Mean *d* is the observed *d* across studies, corrected *d* is corrected for measurement reliability in job performance, *K* is the number of coefficients, *N* is the number of participants, Var. is the variance of the estimate of the corrected *d*, CRI is an 80% credibility interval, and % Art. is the percentage of variance due to the artifacts of sampling error and measurement unreliability.

In our database, industrial samples were associated with a mean corrected *d* of $-.12$ ($K = 41$, $N = 36,326$), variance was $.013$, and the credibility interval ranged from $-.27$ to $.02$. Similarly, public-sector and educational jobs were associated with a mean corrected *d* of $-.20$ ($K = 7$, $N = 769$) and a variance of $.019$ when a large sample was removed. In contrast, military and police jobs were associated with a mean corrected *d* of $.00$ ($K = 12$, $N = 8,260$) and a mean corrected *d* of $.07$ ($K = 11$, $N = 5,271$) when one large sample was removed. Variability was relatively high (variance was $.086$, or $.123$ when the large sample was removed), approximately 3.5 to 10 times as much variance as in other studies. Overall, military studies may have a different pattern of means and variances than do other job categories.

Table 2 also presents our results for date of publication. We grouped studies into three time periods that were approximately equal in terms of the number of studies. Studies from

1969 to 1989 were associated with a mean corrected $d = -.15$. The mean d for the period from 1989 to 2001 was $-.12$, while the mean d for studies from 2002 to 2009 was $-.05$. Thus, there is some evidence that male and female differences in performance evaluations may be decreasing, although credibility intervals do overlap. A similar pattern was obtained across the three time periods when large samples were removed. A reviewer suggested that this pattern of results might be consistent with an increasing number of capable females entering the workforce as time passed.

Results in Table 2 suggest little in the way of moderation for publication status. The corrected mean d for data from journals is $-.11$, while the mean d associated with unpublished sources (i.e., dissertations, theses, and convention presentations) is $-.12$.

There is no clear trend regarding the moderator of job complexity (again, see Table 2). Lower complexity jobs were associated with a mean corrected d of $-.07$. Medium complexity jobs and medium-high complexity jobs were both associated with a corrected d of $-.13$ (and removing two large medium-high complexity samples resulted in a d of $-.07$).

Results in Table 2 also do not suggest a clear pattern of moderation for administrative versus research measures of performance. Administrative performance differences are associated with a mean d of $-.07$, while research measures are associated with a mean d of $-.11$. However, the pattern reverses itself when large samples are removed. For publication status, job complexity, and research purpose moderators, the credibility intervals overlap considerably.

Discussion

There are multiple views in the literature regarding the magnitude (and direction) of gender differences in organizational measures of job performance. Some research suggests that males may, on average, receive higher performance ratings than females receive (e.g., Nieva & Gutek, 1980, as well as Greenhaus & Parasuraman, 1993). Yet, EST suggests that specific status cues regarding job-relevant information (which can also be conceptualized as individuating information) can be more important than diffuse status cues in judgments of job performance. Further, the gender similarities hypothesis (Hyde, 2005) stresses the lack of differences on many psychological individual difference variables that act as theoretical antecedents of job performance.

Overall analyses. We conducted a meta-analysis to investigate gender-based differences on measures of job performance in field studies. In contrast to some previous suggestions (e.g., Nieva & Gutek, 1980), we found that mean gender differences in measures of job performance were generally what one might consider small (i.e., standardized differences smaller than approximately .20 in magnitude; cf. Murphy & Myors, 2004, for a discussion of effect sizes). That is, mean performance levels for males and females are generally more similar than different. If anything, on average, females tended to slightly outscore males. Thus, consistent with EST, data from our meta-analysis support gender similarity for the variable of job performance.

To our knowledge, the current analyses are the most extensive and clearest meta-analytic estimates of gender differences available when on-the-job performance is the focal construct. That is, other meta-analyses examined constructs such as hiring suitability in laboratory studies (e.g., the overall mean d of .41 from Olian et al., 1988). Another meta-analysis focused on the construct of leader performance rather than job performance (Eagly et al., 1995). This latter study also aggregated across laboratory studies, field studies, and a heterogeneous mix of dependent measures (e.g., performance ratings by others, performance ratings by self, satisfaction with the leader, etc.). Thus, we conducted our analysis to provide a clearer picture of gender differences on the construct of operational job performance.

Our results are also different from those of Sackett et al.'s (1991) large primary study in which $d = .07$ (for many blue-collar jobs).⁵ Our more generalizable, meta-analytic difference of $-.11$ is in a different direction (i.e., has a different sign), suggesting that if there is a difference in performance, females generally receive higher performance ratings and/or perform higher. Further, our study cumulates results across approximately 60 samples from a variety of types of jobs and samples.

Also, given that many of our meta-analytic d s were negative, it appears that using performance ratings as a complete or partial predictor of future organizational behavior is unlikely to lead to adverse impact for females (i.e., hiring a smaller proportion of women) in many jobs. However, this may not necessarily be the case for some jobs with a male stereotype (Davison & Burke, 2000).

Overall, our data suggest that males' and females' mean measured levels of job performance are highly similar, and *if* there is a difference, females have a slight advantage. Thus, Hypothesis 1 was supported.

Promotion potential analyses. Our analyses of promotion potential suggest that females are rated lower than males. That is, for the six matched studies, females had slightly better current job performance ratings than did males ($d = -.09$) but *at the same time* had slightly lower promotability ratings ($d = .12$). Thus, Hypothesis 2 was supported.

Taken together, our contrasting results for Hypothesis 1 and Hypothesis 2 are interesting. First, we note that these results are consistent with EST; that is, specific cues (individuating information) are more readily available when rating current performance, but when rating promotion potential, the lack of specific information gives diffuse cues relatively more room to operate. These results could also be interpreted as supportive of the notion in role congruity theory that posits that females may have less access to higher level jobs via the promotion process (Eagly & Karau, 2002). Our empirical results could also be seen as consistent with the lack-of-fit model (Heilman, 1983; Lyness & Heilman, 2006) and concerns that females are perceived by managers as having more work-life conflict (Hoobler, Wayne, & Lemmon, 2009). Again, the increased ambiguity for predicting performance on a new job may be associated with increased use of stereotypes or diffuse cues versus ratings of performance on an existing job (see also Eagly et al., 1995). We urge substantial future research targeted at this finding. We suggest that the dynamic of greater ambiguity of predicting future job performance may also occur during the personnel hiring or selection process (more on this below).

Regarding the concept of gender bias, the number of negative (and near-zero) meta-analytic *ds* suggests that relatively straightforward theoretical views (i.e., bias overpowers qualifications) do not necessarily apply to the assessment of current, operational job performance. Similarly, although our managerial sample was small, female managers were not rated lower on performance than male managers. Thus, our job performance results are not generally supportive of role incongruity theories for the variable of job performance (as per Eagly & Karau, 2002), although we note that our managers were more middle-level managers than upper-level managers. Again, a theoretical explanation for our findings is that specific status cues (i.e., job-related information) are present when supervisors rate subordinates. Other individuals might call this individuating information (e.g., Eagly et al., 1995; Eagly et al., 1992; Reskin, 2000).

Our results do not directly address more complex and subtle theories, such as double standards in which (a) there are higher standards invoked for lower status individuals to get a high evaluation (i.e., relatively easier standards are used for higher status individuals to get a high evaluation) and (b) at the same time, lower standards are invoked for lower status individuals when giving evaluations of minimal competence (i.e., relatively harder standards are used for higher status individuals to be deemed minimally competent; Biernat & Manis, 1994; see also Maurer & Taylor, 1994). Such theories deserve more theoretical and empirical attention. Similarly, the delineation of staff versus line jobs (and associated beliefs and stereotypes) may also be an important distinction in understanding performance assessment and bias (e.g., Lyness & Heilman, 2006).

A reviewer also suggested the possibility of bias against females despite the reported similar job performance ratings. That is, females have some noted advantages in precursors of job performance, yet the *ds* for job performance ratings are close to zero. In particular, we noted that females tend to score higher in a series of samples on measures of declarative job knowledge in one military study and that females, on average, have many similar levels of job-related personality traits. Further, females tend to have higher levels of social skills, as assessed by self-report measures of samples of college students. Studies by Riggio and colleagues (e.g., Riggio, 1986; Riggio, Tucker, & Coffaro, 1989) report *ds* of $-.23$ and $-.46$ in two samples. Some meta-analytic work suggests the advantage by females is even higher in terms of "reading" nonverbal social information, as *ds* sometimes approach -1.0 when a wide variety of verbal and nonverbal information is available to individuals interpreting social skills (Hall, 1978; see also Hall, Carter, & Horgan, 2000). In sum, the reviewer noted that, although *ds* are relatively small in magnitude, bias might exist against females because the bias decreases *ds* to their current levels that ought to be higher given trends in antecedents of job performance.

We suggest future research on this complex interplay of variables. Such research could lead to important insights, although it would likely require a well-specified model of job performance for a given job. Such models would require substantial care, as a complex array of gender-related abilities might be related to job performance (e.g., spatial ability, muscular endurance, agreeableness or dependability; cf. Hough et al., 2001). And researchers should note that job performance *ds* will not necessarily be as large as precursor *ds* because of regression effects (however, to the extent that precursor *ds* are positive, one could expect that job performance *ds* are also positive). Further, one would need to attend to the intercorrelations of

abilities to carefully investigate how the antecedents of performance and bias would work in concert to influence measures of job performance.

Limitations

All research efforts have limitations, and we note several. First, although our meta-analytic sample size was $K = 61$ studies (for performance ratings and measures of output), use of moderator analyses led to relatively small numbers of coefficients (K) for some analyses (e.g., promotability).

Second, similar to other researchers (Sackett et al., 1991), we could not readily analyze subdimensions of performance. While we looked for this information, it was generally not possible to accurately analyze particular parts of the task performance domain (e.g., interacting with others vs. technical performance).

Third, we did not have access to “true scores” for job performance. Nor were we able to find studies such that all factors other than gender were controlled. Thus, given our desire to focus on field studies, we did not directly address ratings bias (as might be possible in the laboratory).

Future Research

We suggest that future research address different dimensions of job performance. One approach is to use the gender roles perspective to make predictions about performance differences. This perspective suggests that females are more likely to learn to deal with the world in a social or communal way. Males may have a more agentic approach that emphasizes more task-oriented behaviors (Eagly & Johannesen-Schmidt, 2001). These emphases may give rise to different gender-related skills and different levels of performance on factors such as interpersonal tasks versus technical tasks—which might vary in emphasis across different jobs.

In particular, research might pursue gender differences within work sample measures of job performance. Work samples, as measures of job performance, have the advantage of being more standardized in terms of opportunities for observation and scoring than are many supervisory ratings of job performance. They can also be designed to focus on particular skill sets, so they could foster a more construct-oriented approach to examining gender performance differences. Unfortunately, we found no work samples that were conceptualized as measures of job performance in our gender-based literature search. Similar work might also be applied to assessment centers, and “head-to-head” tests of gender group differences on work samples, assessment centers, ratings of job performance, and ratings of promotability could be conducted.

We also suggest more study of measures of promotion potential. Again, we found only eight such studies. Researchers might also pursue meta-analyses of withdrawal behaviors (e.g., lateness, absenteeism, turnover).

We also encourage work on extra-role behavior or organizational citizenship behaviors (OCBs; Borman, White, Pulakos, & Oppler, 1991; Vey & Campbell, 2004). Future research

could focus on issues such as the context or situation and on the expectations of those evaluating performance. For example, various dimensions of OCBs might be expected more for men or women (e.g., altruism and courtesy dimensions are considered “in role” behavior more for women than for men; cf. Kidder & McLean-Parks, 2001). Likewise, it is possible that withholding altruistic behavior may be more harmful to evaluations of women than of men (Heilman & Chan, 2005).

We also suggest continuing research into gender differences in hiring. Theoretically, hiring might be viewed as similar to the promotion process. That is, the prediction of future performance might allow more diffuse status cues to influence pre-employment-related judgments because there can be less individuating information related to the quality of the applicants. While there has been substantial work on gender differences on objective tests of variables such as cognitive ability (e.g., Hough et al., 2001), there is less research on more subjective predictors of job performance such as application blanks, interviews, ratings in work sample tests, and so on. Again, we suggest such work use field samples of job applicants whenever possible.

It would also be useful to conduct more research on multiple, potentially interacting, demographic variables. For example, as implied by a reviewer, it would be interesting to investigate if gender and ethnicity combine to influence performance assessments. Similarly, the joint role of age and gender, or age and ethnicity, on job performance might reveal how diffuse status cues interrelate to stereotypes. Such research designs could be used to study other dependent variables such as hiring, promotion, and salary.

Finally, we suggest research into gender differences in political skills. One researcher found that although females performed better on the job than males did, the females were rated as having less influence and less centrality to the dominant coalition of key decision makers (Brass, 1985). This could limit or influence perceptions of promotion potential. Other researchers have suggested that females may have lower levels of organizational political skills, thereby reducing the opportunity to interact with key decision makers and decreasing their chances of promotion (Ferris, Frink, Galang, Zhou, & Howard, 1996; Perrewé & Nelson, 2004).

In sum, we conducted a meta-analysis of gender differences in job performance in field studies. Our findings are generally consistent with predictions of small differences. Females scored slightly higher than males, although credibility intervals often included zero. On the other hand, ratings of promotion potential tended to favor males. Thus, we suggest that ratings of promotability deserve further research attention, that results are consistent with EST, and that future research should consider more complex models regarding any gender bias and job performance differences. We look forward to the knowledge and continued insights that such research will bring.

Notes

1. We used the *d* statistic (i.e., a standardized, mean gender group difference) to provide an index of group differences. The numerator of the *d* statistic is the mean male rating minus the mean female rating. The denominator is the pooled or averaged value of the male and female group standard deviations. Thus, a *d* of $-.25$ would indicate that the mean male score is one quarter of an averaged standard deviation lower than the mean female score.

2. There was one influential case in coding effect sizes in which one coder mistakenly reversed the sign on the effect size. We trimmed this case (Lipsey & Wilson, 2001) from our analyses of agreement. Interrater agreement was .59 with the case and 1.00 without the case.

3. One study reported both an objective and a subjective measure of job performance. We used the subjective measure in overall analyses, given its similarity to most other measures of job performance in this analysis.

4. There were several large samples ($N > 2,000$) that might have influenced results in this analysis (and throughout the rest of the article). In the text, we generally discuss results for all samples unless the large samples have a marked influence on results.

5. We were unable to incorporate Sackett, Dubois, and Noe's (1991) data into our analysis, given the summary form (e.g., frequency distribution) in which it was reported.

References

Asterisks denote studies used in our meta-analysis.

- *Abraham, J. D., & Morrison, J. D. 2003. Relationship between the performance perspectives inventory's implementing scale and job performance of incumbents in a dual security/service role. *Applied HRM Research*, 8(2): 73-76.
- *Arvey, R. D., Miller, H. E., Gould, R., & Burch, P. 1987. Interview validity for selecting sales clerks. *Personnel Psychology*, 40: 1-12.
- *Barrick, M. R., Mount, M. D., & Strauss, J. P. 1994. Antecedents of involuntary turnover due to a reduction in force. *Personnel Psychology*, 47: 515-535.
- Bartlett, C. J., Bobko, P., Mosier, S. B., & Hannan, R. 1978. Testing fairness with a moderated multiple regression strategy: An alternative to differential analysis. *Personnel Psychology*, 31: 233-241.
- Berger, J., Rosenholtz, S. J., & Zelditch, M. 1980. Status organizing processes. *Annual Review of Sociology*, 6: 479-508.
- Berger, J., Wagner, D. G., & Zelditch, M. 1985. Introduction: Expectation states theory. In J. Berger & M. Zelditch (Eds.), *Status, rewards, and influence*: 1-72. San Francisco: Jossey-Bass.
- Biernat, M., & Manis, M. 1994. Shifting standards and stereotype based judgments. *Journal of Personality and Social Psychology*, 66: 5-20.
- Borman, W. C., White, L. A., Pulakos, E. D., & Oppler, S. H. 1991. Models of supervisory job performance ratings. *Journal of Applied Psychology*, 76: 863-872.
- Bowen, C. C., Swim, J. K., & Jacobs, R. R. 2000. Evaluating gender biases on actual job performance of real people: A meta-analysis. *Journal of Applied Social Psychology*, 30: 2194-2215.
- *Brass, D. J. 1985. Men's and women's networks: A study of interaction patterns and influence in an organization. *Academy of Management Journal*, 28: 327-343.
- *Byron, K. 2007. Male and female managers' ability to "read" emotions: Relationships with supervisor's performance ratings and subordinates' satisfaction ratings. *Journal of Organizational and Occupational Psychology*, 80: 713-717.
- Campbell, J. P., Gasser, M. B., & Oswald, F. L. 1996. The substantive nature of job performance variability. In K. R. Murphy (Ed.), *Individual differences and behavior in organizations*: 258-299. Hillsdale, NJ: Lawrence Erlbaum.
- Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. 1993. A theory of performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations*: 35-70. San Francisco: Jossey-Bass.
- Cascio, W. F. 1997. *Applied psychology in human resource management*. Upper Saddle River, NJ: Prentice-Hall.
- Cleary, T. A. 1968. Test bias: Prediction of grades of negro and white students in integrated colleges. *Journal of Educational Measurement*, 5: 115-124.
- Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C. O. L. H., & Ng, K. Y. 2001. Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology*, 86: 425-445.
- *Confidential Industry Technical Report.
- *Confidential Industry Technical Report.
- *Confidential Industry Technical Report.
- *Cook, C. D. 1984. *The validation of an assessment center procedure for the early identification of supervisory job performance and potential for advancement*. Unpublished doctoral dissertation, University of Wisconsin-Madison.

- Correll, S. J., & Ridgeway, C. L. 2003. Expectation states theory. In J. Delamater (Ed.), *Handbook of social psychology*: 29-51. New York: Kluwer Academic/Plenum.
- *Crane, J. B. 1981. *The interrelationship of self-concept and performance and selected characteristics of extension 4-H youth agents within the Mississippi cooperative extension service*. Unpublished doctoral dissertation, Mississippi State.
- *Cropanzano, R., Rupp, D. E., & Zinta, S. B. 2003. The relationship of emotional exhaustion to work attitudes, job performance, and organizational citizenship behaviors. *Journal of Applied Psychology*, 88: 160-169.
- Davison, H. K., & Burke, M. J. 2000. Sex discrimination in simulated employment contexts: A meta-analytic investigation. *Journal of Vocational Behavior*, 56: 225-248.
- *Dereinda, R. 1984. *The relationships between job satisfaction and job performance among county agents in the Wisconsin cooperative extension service*. Unpublished doctoral dissertation, University of Wisconsin-Madison.
- *Diefendorff, J. M., Brown, D. J., Kamin, A. M., & Lord, R. G. 2002. Examining the roles of job involvement and work centrality in predicting organizational citizenship behaviors and job performance. *Journal of Organizational Behavior*, 23: 93-108.
- Dovidio, J. F., Ellyson, S. L., & Keating, C. F. 1988. The relationship of social power to visual displays of dominance between men and women. *Journal of Personality and Social Psychology*, 54: 233-242.
- Dunbar, S. R., & Novick, M. R. 1988. On predicting success in training for men and women: Examples from the Marine Corps clerical specialties. *Journal of Applied Psychology*, 73: 545-550.
- Eagly, A. H., & Johannesen-Schmidt, M. C. 2001. The leadership styles of women and men. *Journal of Social Issues*, 57: 781-797.
- Eagly, A. H., & Karau, S. J. 2002. Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109: 573-598.
- Eagly, A. H., Karau, S. J., & Makhijani, M. G. 1995. Gender and effectiveness of leaders: A meta-analysis. *Psychological Bulletin*, 117: 125-145.
- Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. 1992. Gender and evaluation of leaders: A meta-analysis. *Psychological Bulletin*, 111: 3-22.
- Farh, J., & Dobbins, G. H. 1989. Effects of self-esteem on leniency bias in self-reports of performance: A structural equation model analysis. *Personnel Psychology*, 42: 835-850.
- *Ferris, G. R., Bergin, T. G., & Wayne, S. J. 1988. Personal characteristics, job performance, and absenteeism of public school teachers. *Journal of Applied Social Psychology*, 18: 552-563.
- Ferris, G. R., Frink, D. D., & Galang, M. C., Zhou, J., & Howard, J. L. 1996. Perceptions of organizational politics: Predictions, stress related implications, and outcomes. *Human Relations*, 49: 233-266.
- Freese, L., & Cohen, B. P. 1973. Eliminating status generalization. *Sociometry*, 36: 177-193.
- *Glascock, L. B. 1995. *Attributional style as a predictor of work performance*. Unpublished doctoral dissertation, George Washington University.
- *Green, C., Jegadeesh, N., & Tang, Y. 2009. Gender and job performance: Evidence from Wall Street. *Financial Analysts Journal*, 65: 1-14.
- Greenhaus, J. H., & Parasuraman, S. 1993. Job performance attributions and career advancement prospects: An examination of gender and race effects. *Organizational Behavior and Human Decision Processes*, 55: 273-287.
- *Greenhaus, J. H., Parasuraman, S., & Wormley, W. M. 1990. Effects of race on organizational experiences, job performance evaluations, and career outcomes. *Academy of Management Journal*, 33: 64-86.
- Hall, J. A. 1978. Gender differences in decoding nonverbal cues. *Psychological Bulletin*, 85: 845-857.
- Hall, J. A., Carter, J. D., & Horgan, T. G. 2000. Gender differences in nonverbal communication of emotion. In A. Fischer (Ed.), *Gender and emotion: Social psychological perspectives*: 97-117. Cambridge, UK: Cambridge University Press.
- *Harris, K. J., Kacmar, K. M., & Carlson, D. S. 2006. An examination of temporal variables and relationship quality on promotability ratings. *Group and Organization Management*, 31: 677-699.
- Harrison, D., Newman, D., & Roth, P. L. 2006. How important are job attitudes? Meta-analytic comparisons of integrative behavioral outcomes and time sequences. *Academy of Management Journal*, 49: 305-326.
- Harrison, D. A., & Shaffer, M. A. 1994. Comparative examinations of self-reports and perceived absenteeism norms: Wading through Lake Wobegon. *Journal of Applied Psychology*, 79: 240-251.

- Heilman, M. E. 1983. Sex bias in work settings: The lack of fit model. *Research in Organizational Behavior*, 5: 269-298.
- Heilman, M. E. 2001. Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues*, 57: 657-674.
- Heilman, M. E., & Chen, J. J. 2005. Same behavior, different consequences: Reactions to men's and women's altruistic citizenship behavior. *Journal of Applied Psychology*, 90: 431-441.
- Heilman, M. E., & Haynes, M. C. 2005. No credit where credit is due: Attributional rationalization of women's success in male-female teams. *Journal of Applied Psychology*, 90: 905-916.
- *Helmick, T. A. 1987. *The effect of gender, job classification and self-esteem upon job performance*. Unpublished master's thesis, Central Missouri State University.
- *Hoobler, J. M., Wayne, S., & Lemmon, G. 2009. Bosses' perceptions of family-work conflict and women's promotability: Glass ceiling effects. *Academy of Management Journal*, 52: 939-957.
- Hough, L. M., Oswald, F. L., & Ployhart, R. E. 2001. Determinants, detection and amelioration of adverse impact in personnel selection procedures: Issues, evidence and lessons learned. *International Journal of Selection and Assessment*, 9: 152-194.
- Hunter, J. E. 1983a. A causal analysis of cognitive ability, job knowledge, and job performance, and supervisory ratings. In R. Landy, S. Zedeck, & J. Cleveland (Eds.), *Performance measurement and theory*: 257-266. Hillsdale, NJ: Lawrence Erlbaum.
- Hunter, J. E. 1983b. *Test validation for 13,000 jobs: An application of job classification and validity generalization for the General Aptitude Test Battery (GATB)*. USES Test Research Rep. No. 450. Washington, DC: U.S. Department of Labor.
- Hunter, J. E., & Schmidt, F. L. 2004. *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA: Sage.
- Hyde, J. S. 2005. The gender similarities hypothesis. *American Psychologist*, 60: 581-592.
- *Igbaria, M., & Baroudi, J. J. 1995. The impact of job performance evaluations on career advancement prospects: An examination of gender differences in the IS workplace. *MIS Quarterly*, 19(1): 107-123.
- *Jepsen, D. M., & Rodwell, J. J. 2009. Justice in the workplace: The centrality of social versus judgmental predictors of performance varies by gender. *International Journal of Human Resource Management*, 20: 2066-2083.
- Jagacinski, C. M. 1987. Androgyny in a male-dominated field: The relationship of sex-typed traits to performance and satisfaction in engineering. *Sex Roles*, 17: 529-547.
- *Joerger, J. H. 1984. *A field study of fear of success, internal-external locus of control and performance ratings among professional men and women*. Unpublished doctoral dissertation, University of Northern Colorado.
- *Joshi, A., Liao, J., & Jackson, S. 2006. Cross-level effects of workplace diversity on sales performance and pay. *Academy of Management Journal*, 49: 459-481.
- Kidder, D. L., & McLean Parks, J. 2001. The good soldier: Who is s(he)? *Journal of Organizational Behavior*, 22: 939-959.
- King, E. B., Hebl, M. R., George, J. M., & Matusik, S. F. (2010). Understanding tokenism: Antecedents and consequences of a psychological climate of gender inequity. *Journal of Management*, 36: 482-510.
- Landy, F. J., & Farr, J. L. 1980. Performance rating. *Psychological Bulletin*, 87: 72-107.
- Lipsey, M. W., & Wilson, D. B. 2001. *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- *Lyness, K. S., & Heilman, M. E. 2006. When fit is fundamental: Performance evaluations and promotions of upper-level female and male managers. *Journal of Applied Psychology*, 91: 777-783.
- MacGillivray, E. D., Beecher, H. J. M., & Golden, D. 2009. Legal developments—Gender issues now, and increasingly in the US and EU spotlight. *Global Business and Organizational Excellence*, 28(3): 79-86.
- *Marion-Landais, C. A. 1999. *Dyadic agreement about ideal leadership behaviors and its relationship to job satisfaction and performance ratings*. Unpublished doctoral dissertation, University of South Florida.
- Maurer, T. M., & Taylor, M. A. (1994). Is sex by itself enough: An exploration of gender bias issues in performance appraisal. *Organizational Behavior and Human Decision Processes*, 60: 231-251.
- McCloy, R. A., Campbell, J. P., & Cudeck, R. (1994). A confirmatory test of a model of performance determinants. *Journal of Applied Psychology*, 79: 493-505.
- McKay, P. F. 2009. Perspectives on adverse impact in work performance: What we know and what we could learn more about. In J. Outtz (Ed.), *Adverse impact: Implications for organizational staffing and high stakes selection*: 249-270. New York: Routledge.

- McKay, P. F., & McDaniel, M. A. 2006. A reexamination of Black–White mean differences in work performance: More data, more moderators. *Journal of Applied Psychology*, 91: 538–554.
- *Michael, D. F. 2004. *Gender differences in information-exchange and outcomes at work*. Unpublished doctoral dissertation, Auburn University.
- *Mills, A. 1990. *Predicting police performance for differing gender and ethnic groups: A longitudinal study*. Unpublished doctoral dissertation, California School of Professional Psychology.
- Mintzberg, H. 2001. The yin and yang of managing. *Organizational Dynamics*, 29: 306–312.
- *Mobley, W. H. 1982. Supervisor and employee race and sex effects on performance appraisals: A field study of adverse impact and generalizability. *Academy of Management Journal*, 25: 598–606.
- *Morrison, J. D., & Abraham, J. D. 2003. Relationship between the performance perspectives inventory's selling scale and job performance of real estate agents. *Applied HRM Research*, 8(2): 85–88.
- *Motowildo, S. J., & Schmit, M. J. 1997. *Performance assessment and interview procedures for store manager and store associate positions* (Final tech. rep.). University of Florida, Human Resource Research Center.
- Murphy, K. R., & Myers, B. 2004. *Statistical power analysis: A simple and general model for traditional and modern hypothesis tests* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.
- Ng, T. W. H., Eby, L., Sorensen, K. L., & Feldman, D. C. 2005. Predictors of objective and subjective career success. A meta-analysis. *Personnel Psychology*, 58: 367–408.
- Nieva, V. F., & Gutek, B. A. 1980. Sex effects on evaluation. *Academy of Management Review*, 5: 267–276.
- *O'Connell, M. S., Hartman, N. S., McDaniel, M. A., Grubb, W. L., & Lawrence, A. 2007. Incremental validity of situational judgment tests for task and contextual performance. *International Journal of Selection and Assessment*, 15: 19–29.
- Olian, J. D., Schwab, D. P., & Haberfield, Y. 1988. The impact of applicant gender compared to qualifications on hiring recommendations: A meta-analysis of experimental studies. *Organizational Behavior and Human Decision Processes*, 41: 180–195.
- *Ostroff, C., Atwater, L. E., & Feinberg, B. J. 2004. Understanding self–other agreement: A look at rater and rater characteristics, context, and outcomes. *Personnel Psychology*, 57: 333–375.
- *Patterson, T. F., Jr. 1984. *A study to determine the relationship between rural New England extension agent educational orientation and job performance*. Unpublished doctoral dissertation, Indiana University.
- *Pazy, A., & Oron, I. 2001. Sex proportion and performance evaluation among high-ranking military officers. *Journal of Organizational Behavior*, 22: 689–702.
- Perrewe, P. L., & Nelson, D. L. 2004. The facilitative role of political skill. *Organizational Dynamics*, 33: 366–378.
- *Phillips, J. F. 1992. Predicting sales skills. *Journal of Business and Psychology*, 7: 151–160.
- *Phillips, J. F. 1993. Predicting negotiation skills. *Journal of Business and Psychology*, 7: 403–411.
- *Plamondon, K. E., & Schmitt, N. 2000. *Validity and subgroup differences of combinations of predictors as a function of research design*. Paper presented at the meeting of the Society for Industrial and Organizational Psychology, New Orleans.
- *Ployhart, R. E., Wiechmann, D., Schmitt, N., Sacco, J. M., & Rogg, K. 2003. The cross-cultural equivalence of job performance ratings. *Human Performance*, 16: 49–79.
- *Pulakos, E. D., Schmitt, N., & Chan, D. 1996. Models of job performance ratings: An examination of rate race, rate gender, and rater level effects. *Human Performance*, 9: 103–119.
- *Pulakos, E. D., & Wexley, K. N. 1983. The relationship among perceptual similarity, sex, and performance ratings in manager-subordinate dyads. *Academy of Management Journal*, 26: 129–139.
- *Pulakos, E. D., White, L. A., Oppler, S. H., & Borman, W. C. 1989. Examination of race and sex effects on performance ratings. *Journal of Applied Psychology*, 74: 770–780.
- Reskin, B. F. 2000. The proximate causes of employment discrimination. *Contemporary Sociology*, 29: 319–328.
- Riggio, R. E. (1986). Assessment of basic social skills. *Journal of Personality and Social Psychology*, 51: 649–660.
- Riggio, R. E., Tucker, J., & Coffaro, D. 1989. Social skills and empathy. *Personality and Individual Differences*, 11: 93–99.
- Robertson, T. M. 1986. Fundamental strategies for wage and salary administration. *Personnel Journal*, 65(11): 120–132.
- Roth, P. L., Huffcutt, A. I., & Bobko, P. 2003. Ethnic group differences in measures of job performance: A new meta-analysis. *Journal of Applied Psychology*, 88: 725–740.

- *Saad, S., & Sackett, P. R. 2002. Investigating differential prediction by gender in employment-oriented personality measures. *Journal of Applied Psychology*, 87: 667-674.
- Sackett, P. R., Dubois, C. L. Z., & Noe, A. W. 1991. Tokenism in performance evaluation: The effects of work group representation on male-female and White-Black differences in performance ratings. *Journal of Applied Psychology*, 76: 263-267.
- Schein, V. E. 2001. A global look at psychological barriers to women's progress in management. *Journal of Social Issues*, 57: 675-688.
- *Schmidli, K. R. 1969. *Effectiveness of teachers of the educable mentally retarded as related to preparation*. Unpublished doctoral dissertation, Colorado State College.
- Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. 1986. The impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology*, 71: 432-439.
- Schmidt, F. L., Hunter, J. E., Outerbridge, A. N., & Goff, S. 1988. Joint relation of experience and ability with job performance: Test of three hypotheses. *Journal of Applied Psychology*, 73: 46-57.
- Schmidt, F. L., & Le, H. 2004. *Software for the Hunter-Schmidt meta-analysis methods*. Department of Management and Organization, University of Iowa, Iowa City.
- *Schmit, M. J., & Motowildo, S. J. 1995. *Development and validation of a situational interview and personality test for selecting sales associates* (Final tech. rep.). University of Florida, Human Resource Research Center.
- Schoorman, F. D., Mayer, R. C., & Davis, J. H. 2007. An integrative model of organizational trust: Past, present, and future. *Academy of Management Review*, 32: 344-354.
- *Shore, L. M., & Thornton, G. C. 1986. Effects of gender on self- and supervisory ratings. *Academy of Management Journal*, 29: 115-129.
- *Shore, T. H. 1992. Subtle gender bias in the assessment of managerial potential. *Sex Roles*, 27: 499-515.
- Swim, J., Borgida, E., Maruyama, G., & Myers, D. G. 1989. Joan McKay versus John McKay: Go gender stereotypes bias evaluations. *Psychological Bulletin*, 105: 409-429.
- *Thacker, R. A., & Wayne, S. J. 1995. An examination of the relationship between upward influence tactics and assessments of promotability. *Journal of Management*, 21: 739-756.
- *Tsui, A. S., & Gutek, B. A. 1984. A role set analysis of gender differences in performance, affective relationships, and career success of industrial middle managers. *Academy of Management Journal*, 27: 619-635.
- Tsui, A. S., & O'Reilly, C. A. 1989. Beyond simple demographic effects: The importance of relational demography in superior-subordinate dyads. *Academy of Management Journal*, 32: 402-423.
- Vey, M. A., & Campbell, J. P. 2004. In-role or extra-role organizational citizenship behavior: Which are we measuring? *Human Performance*, 17: 119-135.
- Viswesvaran, C., Ones, D. S., & Schmidt, F. L. 1996. Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology*, 81: 557-574.
- Viswesvaran, C., Schmidt, F. L., & Ones, D. S. 2005. Is there a general factor in ratings of job performance? A meta-analytic framework for disentangling substantive and error influences. *Journal of Applied Psychology*, 90: 108-131.
- *Wanous, J. P., Stumpf, S. A., & Bedroisan, H. 1979. Job survival of new employees. *Personnel Psychology*, 32: 651-662.
- *Weekley, J. A., & Jones, C. 1999. Further studies of situational tests. *Personnel Psychology*, 52: 679-700.
- *Wexley, K. N., & Pulakos, E. D. 1982. Sex effects on performance ratings in manager-subordinate dyads: A field study. *Journal of Applied Psychology*, 67: 433-439.
- Wood, J. 2008. Methodology for dealing with duplicate effects in a meta-analysis. *Organizational Research Methods*, 11: 79-95.
- Wood, W., & Karten, S. J. 1986. Sex difference in interaction style as a product of perceived differences in competence. *Journal of Personality and Social Psychology*, 50: 341-347.

WHEN CAN WOMEN CLOSE THE GAP? A META-ANALYTIC TEST OF SEX DIFFERENCES IN PERFORMANCE AND REWARDS

APARNA JOSHI

Pennsylvania State University

JOOYEON SON

University of Illinois, Urbana-Champaign

HYUNTAK ROH

Yonsei University

Drawing on macro and micro domains in gender research, we meta-analytically test whether occupation-, industry-, and job-level factors mitigate or exacerbate differences in performance evaluations ($k = 93$; $n = 95,882$) and rewards ($k = 97$; $n = 378,850$) between men and women. Based on studies conducted across a variety of work settings and spanning nearly 30 years, we found that the sex differences in rewards ($d = .56$) (including salary, bonuses, and promotions) were 14 times larger than sex differences in performance evaluations ($d = .04$), and that differences in performance evaluations did not explain reward differences between men and women. The percentage of men in an occupation and the complexity of jobs performed by employees enhanced the male–female gap in performance and rewards. In highly prestigious occupations, women performed equally, but were rewarded significantly lower than men. Only a higher representation of female executives at the industry level enabled women to reverse the gender gap in rewards and performance evaluations. Our configurational analysis also revealed that some occupation-, industry-, and job-level attributes of the work context jointly contributed to differences in rewards and performance evaluations.

In 1982, when Ann Hopkins was denied partnership at a prestigious accounting firm despite a stellar performance record, her attorneys presented compelling evidence before the Supreme Court of the United States that the decision was a direct outcome of gender-based stereotyping and discrimination in the firm (Fiske, Bersoff, Borgida, Deaux, & Heilman, 1991; see also *Price Waterhouse v. Hopkins* 109 S. Ct. 1775 (1989)). Such instances of overt stereotyping and egregious gender discrimination are rare in workplaces today. Nevertheless, more than two decades since the historic Supreme Court ruling in favor of Ann Hopkins, women continue to remain grossly underrepresented at the

highest levels in organizations, and in many settings receive significantly lower pay and fewer promotions than men (Catalyst, 2008). Indeed, scholars recognize that detecting or remedying gender-based stereotyping, bias, or discrimination continues to pose a substantial challenge in organizations (e.g., Blau & Kahn, 2007; Kalev, Dobbin, & Kelly, 2006; Reskin, 2003). Litigation pressures and the diffusion of symbolic diversity management practices obscure bias and stereotyping, making it unlikely that purposive discrimination attributable to specific organizational agents can be easily detected (Ortiz & Roscigno, 2009; Tilly, 1998). Therefore, rather than focusing research on detecting the (often unmeasurable) psychological processes that underlie gender discrimination at work, some scholars have noted that identifying objective (often measurable) attributes of the workplace context associated with sex differences in employment outcomes offers one important avenue for building an actionable theory of gender inequality (Baron & Pfeffer, 1994; Reskin, 2003).

We thank Jason Colquitt and three anonymous reviewers for their valuable feedback throughout the review process. We are also very grateful to Vilmos Misangyi and Daniel Newman for their helpful feedback on previous versions of the manuscript. Funding for this research was provided by a National Science Foundation Grant (#1307353).

As a step in this direction, we develop and test a theoretical framework that draws on the cumulative research evidence, spanning nearly three decades, on sex differences in organizational rewards and performance evaluations. This evidence points out that although sex differences in organizational rewards and performance evaluations are fairly widespread, the gap between women and men also varies considerably across work contexts. Certainly, in many settings, men outperform women and receive higher rewards, such as pay and promotions (e.g., Greenhaus & Parasuraman, 1993; Sackett, DuBois, & Noe, 1991). However, in some situations, women do receive more favorable performance evaluations and are also able to close the gap in organizational rewards (e.g., Dencker, 2008; Speilerman & Petersen, 1999). The varied nature of the gap between women and men across work contexts offers important directions for further inquiry; it suggests that some contexts may be more susceptible and others more resilient to gender inequality. Building on this premise, our approach seeks to develop a theoretical framework that identifies several structural attributes of the work context that can reinforce or mitigate differences in performance and reward outcomes between men and women.

This framework offers a substantive step forward in gender inequality research for a number of reasons. In both micro (e.g., Cleveland, Vescio, & Barnes-Farrell, 2005) and macro (e.g., DiTomaso, Post, Parks-Yancy, 2007) domains of gender research, there have been calls to integrate across disciplinary silos to better understand the prevalence of inequity at work. Indeed, more than two decades ago, Baron and Pfeffer (1994: 191) noted that:

... missing in most literature on (inequity in organizations) is any attention to the micro-macro connection, the links between social structures, institutions, and organizations on the one hand and, on the other, cognitions, perceptions, interests and behaviors at the individual or small-group level.

To date, management research, although uniquely positioned to make this “micro-macro” connection, has not fully responded to these calls.

We aim at developing a framework that bridges the micro-macro divide in sex differences research by uniting institutional and structural perspectives on gender inequality (e.g., Stainback, Tomaskovic-Devey, & Skaggs, 2010) with sociopsychological perspectives on gender bias (e.g., Eagly, Makhijani, & Klonsky, 1992), to specify how aspects of the work context predict sex differences in performance evaluations

and organizational rewards. Based on these theoretical perspectives, we meta-analytically test the effects of occupation-, industry-, and job-level predictors on sex differences in performance evaluations and reward outcomes among men and women performing comparable jobs. For example, our approach allows us to examine whether the level of complexity of jobs performed by investment bankers explains why female investment bankers earn significantly lower bonuses than male investment bankers, or whether greater parity in performance evaluations among male and female software developers can be attributed to a high proportion of female executives in the information technology (IT) industry. Drawing on multiple strands in gender research, we propose that although gender bias or stereotyping can potentially operate in each of these contexts, factors such as job complexity or representation of female executives at the industry level can strengthen or weaken the extent to which these processes determine employment outcomes for men and women. Based on this framework, we recommend directions for future research on the role that performance evaluation and reward allocation decision making can play in shaping sex differences in rewards or performance within specific contexts.

A second important contribution of this research is that we examine the effects of contextual attributes on both performance evaluations and reward differences between men and women. In order to unpack the sources of gender inequality, it is important to understand whether or not contextual factors that mitigate or enhance sex differences in performance also mitigate or enhance sex differences in pay and promotions (Blau & Kahn, 2007; Cleveland et al., 2005). It is often assumed that psychological processes or structural attributes that drive sex differences in performance evaluations also drive sex differences in outcomes such as pay, promotions, or other organizational rewards. We propose that this assumption needs to be examined more carefully. Consider that, in many contexts, reward allocation decisions (e.g., promotion to equity partnership) often imply zero-sum choices among a pool of qualified candidates, while performance criteria (e.g., billable hours) are not subject to such considerations. By comparing sex differences in performance evaluations versus rewards across various work settings, we hope to shed light on whether various contextual variables have similar effects on these outcomes. Put simply, our framework allows us to test whether in contexts in which men are rewarded at higher levels than women, they also receive higher performance evaluations

than women, and in contexts in which women receive higher evaluations, they also receive higher rewards than men.

On a related note, we also directly test whether performance differences *explain* reward differences among men and women. Research that has examined differences in pay and promotions among men and women rarely account for performance differences (see Castilla, 2008; Blau & Devaro, 2007, as exceptions). Instead, it is often assumed that productivity differences, measured as educational attainment or labor market experience, drive sex differences in rewards (Tomaskovic-Devey, 1993). However, the educational and skill attainments that women have made since the 1980s have surpassed even average workforce-level human capital gains during the period (Bernhardt, Morris, & Handcock, 1995; Blau & Kahn, 2007). These gains should enable women to close any performance gaps with their male counterparts, but the question remains: Do these gains also allow them to close the gender gap in organizational rewards? Our framework integrates past research on sex differences in pay with research on sex differences in performance evaluations to answer this question.

Our meta-analysis is also unique in that we apply a configurational approach to supplement meta-regression procedures and to identify whether *bundles* of occupational, industry, and job attributes jointly contribute to sex differences in performance evaluation and organizational rewards (Fiss, 2011; Ragin, 2008). This approach allows us to identify which contextual factors combine, complement, or act as substitutes for each other in situations in which men outperform, or are rewarded significantly higher than, women. Such an approach also enables us to highlight less obvious combinations of contextual attributes that may be sufficient for observing sex differences in performance evaluations and rewards.

MACRO-MICRO LINKAGES IN GENDER RESEARCH

The proposed framework tests the notion that sex differences in organizational rewards and performance evaluations are a function of “what” the nature of the work is, as well as “where” the work is being done (Reskin, 2003). Reskin (2003: 14) notes that these aspects of “structure and context are fundamental concepts because they highlight the importance of [the] setting on social processes that govern [gender] inequality in the workplace.” Developing this logic, we focus on how job-level (i.e., what the

nature of the work is), as well as occupation- and industry-level (i.e., where the work is being done), attributes can shape sex differences in performance evaluations and rewards in an organization.

Consider that any given work organization operates within a larger industry environment and employs individuals across a diverse array of occupations, performing varied tasks ranging from complex to routine. Recognizing that such an organization (particularly in the United States) also has considerable discretion in designing performance management and compensation practices (Nelson & Bridges, 1999), we propose that specific job, industry, and occupational factors define the overall environment in which organizational agents such as human resources (HR) professionals, executives, and supervisors make evaluative and resource allocation decisions (Dobbin, 2009; Tilly, 1998). This environment comprises specific cultural norms, stereotypic expectations, and status cues associated with gender. This environment also reflects the institutional or market forces that define a work context and influence employment outcomes for men and women. Within this embedding context, we discuss below how reward allocation and performance evaluation practices advantaging men over women can potentially take on an “everyday,” or “business as usual,” quality (Ely & Padavic, 2007; Hultin & Szulkin, 2003; Tilly, 1998).

Occupational Effects on Sex Differences in Performance Evaluations and Rewards

In the past, some research accounts linking occupations to gender inequality have emphasized women’s skill levels and related job choices as antecedents of women’s segregation into lower paying occupations (Bergmann, 1974; Blau & Kahn, 1981). But this research does not fully explain why gender inequality is prevalent despite women’s human capital gains and a steady entry into traditionally male-dominated or higher paying occupations (Gatta & Roos, 2005). We propose that overall prestige and gender typing of an occupation defines normative role expectations, status cues, and stereotypic beliefs about how occupational incumbents are perceived, evaluated, and rewarded, providing important additional explanations for persistent differences between men and women.

Occupational demographic composition. The demographic make-up of an occupation can signal the “appropriateness,” or “fit,” of an occupation for men and for women, driving stereotypic expectations

of role and abilities associated with that occupation (Eagly & Makhijani, 2002; Gorman, 2005; Heilman, 1983). Occupations that are female-dominated (such as nursing or elementary school teaching) may be considered a more appropriate “fit” for women, while occupations that are male-dominated (such as fire-fighting or production engineering) may be considered appropriate for men.

When women enter highly male-dominated occupations, they do not “fit” the stereotypic expectations of abilities expected in that occupation, and therefore experience greater bias and discrimination (Eagly & Makhijani, 2002; Heilman, 1983). Role congruity theory, which has been primarily applied to women in leadership roles, posits that even when female leaders display high levels of performance, their efforts may be discounted by their peers and/or supervisors (e.g., Eagly et al., 1992). This theory recognizes that culturally shared beliefs about the appropriate roles and abilities of men and women in society have widespread effects in the workplace (Eagly & Karau, 1991; Eagly et al., 1992). As such, the contributions of women who occupy roles or display abilities that are atypical relative to established cultural norms tend to be undervalued and discounted at work (Eagly & Karau, 1991; Eagly et al., 1992; Kanter, 1977). These gender-based attributions of skills are often not essential, or even beneficial, for performing various aspects of a job, but can still influence how work is defined and valued in organizations (Ely & Padavic, 2007).

Returning to our opening example, a factor contributing to gender bias in Ann Hopkins’ performance appraisals was her lack of fit in the broad occupational category of professional accountants—an occupational category dominated by men (Fiske et al., 1991). If Ann Hopkins had been an administrative assistant (i.e., in a female-dominated occupational category) in the firm, it is less likely that she would have faced the types of barrier that she encountered. Experimental research shows that women in male-typed settings encounter bias at two levels: First, women are viewed as having less male-typed ability than men (e.g., women are generally viewed as having less agentic leadership ability than men); and second, women who do display male-typed behavior are viewed as less effective than men who display the same behavior (Eagly & Karau, 1991). More recent extensions of this theory have also identified a backlash against female leaders who are perceived as threatening to the male-dominant status quo (Rudman, Moss-Racusin, Phelan, & Nauts, 2012). These perspectives suggest that, in male-dominated occupations, women

may face a “double penalty” when their performance is evaluated, or they are considered for promotions or for bonuses. First, supervisors may view women as less qualified to perform a job or to receive a promotion because their skills or abilities do not fit male-typed norms and expectations. Second, when women do display male-typed behaviors that run counter to female-typed stereotypic expectations, they may be viewed as threatening to the status quo and may be evaluated negatively or denied opportunities for career advancement (Fiske et al., 1991; Rudman et al., 2012).

Across many male-typed occupational contexts, research provides general support for this line of argument. Based on survey data from scientists and engineers across U.S. corporations, research shows that, in these traditionally male-typed occupations, white men received significantly more access to skill development opportunities and higher performance evaluations than any other demographic group, including white women, women of color, or immigrant men (DiTomaso, Post, Smith, Farris & Cordero, 2007). In the context of science and engineering teams, based on “round robin” data gathered from team members, Joshi (2014) found that gender served as an expertise cue such that male scientists received higher expertise evaluations from their male and female colleagues, and, in some circumstances, highly educated women were evaluated *lower* than less educated women by male team members. Men also received more opportunities than women to collaborate in ongoing research projects within teams (Joshi, 2014).

Across other occupational contexts as well, research shows that membership in a demographic group that is dominant in an occupation predicts access to important information such as the availability of training programs and job opportunities (Tomaskovic-Devey & Skaggs, 1999). Cumulatively, this research suggests that the demographic composition of the occupation—specifically, the predominance of men—reflects a cultural context (i.e., gender-related norms, stereotypes, and status cues) that shapes reward allocation and evaluation behavior, with potentially negative employment consequences for women. Thus we posit that:

Hypothesis 1a. The proportion of men in an occupational category enhances sex differences in performance evaluations such that men receive higher performance evaluations than women in occupations with a higher proportion of men.

Hypothesis 1b. The proportion of men in an occupational category enhances sex differences in organizational rewards such that men receive higher rewards than women in occupations with a higher proportion of men.

Occupational prestige. Occupational prestige, the socioeconomic value of an occupation, has been acknowledged as an important variable predicting employment outcomes because it foreshadows occupational mobility and is a proxy for social class (Hodge, Siegel, & Rossi, 1964). The prestige of an occupation may be related to, but also operates independently from, the sex composition of the occupation in predicting income differences between men and women (England, 1979; England & McLaughlin, 1979; Siegel, 1971). Indeed, some highly prestigious occupations, such as dentist or psychologist, are fairly gender-balanced, whereas others, such as astronaut or physician, tend to be male-dominated. Some occupations, such as “automotive salesperson” or “warehouse worker,” are male-dominated, but less prestigious. Not only are occupational prestige hierarchies extremely stable across time, but also there is widespread consensus within societies regarding which occupations are more or less prestigious (Hodge et al., 1964). Because highly prestigious occupations, representing higher social class, require higher investments in human capital (consider lawyers or physicians), we might expect the men and women working in these settings to be similarly qualified and not to experience differential employment outcomes. However, women face several barriers in these contexts as well.

In highly prestigious occupations, barriers to entry and advancement are extremely high. Access to advancement in these settings is controlled via mechanisms such as licensing, formal educational requirements, certification, and sometimes opaque performance and reward criteria (e.g., Weeden, 2002). Research based on social dominance theory suggest that formal and informal practices in these occupational settings constitute a hierarchy-enhancing context (Sidanius & Pratto, 1999; Sidanius, Pratto, Sinclair & van Laar, 1996; Sinclair, Sidanius, & Levin, 2010). “Hierarchy-enhancing agents”—a term broadly applied to social institutions or social ideologies—tend to “support and facilitate the disproportionate allocation of things with positive social value (e.g., high income, prestigious jobs, good education) to dominant social groups and things with negative social value (e.g., prison sentences, poor jobs, inferior education) to subordinate social groups”

(Sidanius, et al., 1996, p. 146). “Hierarchy-attenuating agents,” on the other hand, facilitate an egalitarian allocation of positively valued goods to dominant and subordinate social groups (Sidanius et al., 1996)

Highly prestigious occupations can operate as hierarchy-enhancing agents that favor the distribution of valued goods, such as positive performance appraisals or organizational rewards, to dominant social groups, such as men, over subordinate groups, such as women (Sinclair et al., 2010). Some research shows that, in these types of workplace, key decision makers involved in performance evaluations or wage-setting processes (senior executives, HR professionals, supervisors) are also likely to be high in social dominance orientation—a dispositional attribute that underlies a tendency to favor dominant over subordinate groups (Pratto, Stallworth, Sidanius, Siers, 1997). These dispositional tendencies of key decision makers and the prestige value of the context could serve as mutually reinforcing hierarchy-enhancing factors contributing to inequitable outcomes for women. Moreover, feminist scholars who have commented extensively on the intersections between gender and class note that gender and class relations operate jointly in these employment contexts (Acker, 2006; McCall, 2001). These scholars have argued that wage-setting and supervisory practices function as instruments of control over valued organizational resources, and serve to maintain the dominance of men in high-prestige work contexts that are also avenues for upward class mobility (e.g., Acker, 2006).

In these contexts, even seemingly meritocratic reward allocation and performance evaluation practices can function as hierarchy-enhancing agents by awarding more favorable outcomes to men than to women. For instance, in many prestigious settings, “up or out” promotion policies drive adversarial processes that are primed toward highlighting employees’ weaknesses. In these situations, evaluative decisions may draw on stereotypic information to highlight shortcomings (e.g., “not forceful enough,” “too emotional,” “lacks leadership qualities”) that tend to disadvantage women (Cleveland et al., 2005; Vescio, Gervais, Snyder, & Hoover, 2005). Across prestigious law firms, for example, recent research shows that senior partners making promotion decisions tended to value different types of client strategy across male and female associates: Women were more likely to be promoted when they brought in external clients, but men were promoted when they collaborated with senior partners within the firm (Briscoe & von Nordenflycht, 2014).

Although, on the face of it, both client strategies—internal and external—may contribute to firm revenues, bringing in external clients could be more

challenging and risky, imposing a higher burden on incoming female associates. Perhaps as a reflection of these subtle barriers, several studies in the context of law firms report that promotion rates are significantly lower and turnover rates significantly higher among women than among men (Gorman, 2006; Kay & Hagan, 1998; Spurr & Sueyoshi, 1994). Overall, these findings illustrate how subtle biases might permeate ostensibly meritocratic and gender-neutral performance evaluations and reward allocation practices in prestigious contexts, advantaging men over women (Martin & Myerson, 1998). In less prestigious contexts, barriers to entry and advancement are low, as are the stakes for controlling resources and maintaining the dominant status quo. In these occupational settings, the drive to reward dominant social groups over subordinate groups by means of performance evaluation and reward allocation practices is also likely to be lower, resulting in more equitable outcomes for women (Sinclair et al., 2010). Building on these insights, we propose that:

Hypothesis 2a. The prestige associated with an occupational category enhances sex differences in performance evaluations such that, in highly prestigious occupations, men receive higher performance evaluations than women.

Hypothesis 2b. The prestige associated with an occupational category enhances sex differences in organizational rewards such that, in highly prestigious occupations, men receive higher rewards than women.

Industry Effects on Sex Differences in Performance Evaluations and Rewards

The industry environment is also a distinct and relevant context shaping gender inequality in organizations. The industry context in which a firm operates refers to aspects of the business environment such as the regulatory framework, number and size of competitors, and the level of growth and globalization. Strategy scholars therefore regularly take into account industry attributes such as capital intensity, dynamism, and munificence in predicting firm performance (e.g., McGahan & Porter, 1997; Porter, 1980). The industry setting serves as an embedding context for firms, because firms within an industry compete within similar labor markets and face common institutional pressures that lead to the diffusion of employment practices across these firms (Stainback et al., 2010). Within a single firm encompassing multiple occupational categories, these institutional pressures could potentially influence the

adoption of reward allocation and performance evaluation practices underlying gender inequality.

The institutional context not only frames the adoption and diffusion of employment practices, but also can shape how employees respond to these practices. Research shows that, within an industry, perceptions of discrimination are socially constructed via regulatory and competitive pressures (Hirsch & Kornrich, 2008). Industrial categories are associated with broadly defined norms and expectations regarding antidiscrimination laws (Edelman, Fuller, & Mara-Drita, 2001). In fact, some scholars have found significant between-industry variations in the extent to which employees can “name” a practice as discriminatory, “blame” the employer for that practice, and “claim” redress for the discrimination (see Felstiner, Abel, & Sarat, 1980: 635–636.). Recognizing the importance of the broad industry setting, we propose more specifically that the ascendance of women to managerial and executive levels reflects institutional, labor market, and cultural pressures within an industry, and is an important contingency shaping sex differences in performance evaluations and rewards.

Research finds that the proportion of women in higher ranks of an organization serves as a significant constraint on gender inequality (e.g., Cohen & Broschak, 2013; Cohen, Broschak, & Haveman, 1998; Ely, 1994; Huffman, Cohen, & Pearlman, 2010). Women’s representation in executive positions drives positive outcomes for junior women (Ely, 1994; 1995); senior women also pave the way for greater access to organizational power structures for other women (Huffman et al., 2010; Hultin & Szulkin, 2003). The demographic composition of managerial levels also has been found to reduce the gap in salary and in objective performance outcomes between men and women (Joshi, Liao, & Jackson, 2006; Ostroff & Atwater, 2003). A key logic underlying these findings is that the presence of women in positions of power in organizations has symbolic value—i.e., it signals greater status attainment among women. The presence of women at higher levels also provides more junior women access to career-related advice and mentoring (Ely, 1994; 1995). Finally, the presence of women at higher levels might also imply greater scrutiny of wage-setting, bonus allocation, and performance evaluation practices within organizations (Huffman et al., 2010; Joshi et al., 2006).

Although the research outlined above is focused at the organizational level, the growing role and visibility of industry-level professional women’s networks (e.g., Women on Wall Street, the National Council for Women in Information Technology) suggests that

these arguments will also hold across firms within specific industry categories. These industry-wide network groups often organize conferences and workshops to expose women to visible symbols of success in the industry, and to help them to develop personal strategies to achieve career success. Thus the proportion of women in an industry can serve as a cultural context, shaping status cues and role expectations associated with gender. Moreover, within a specific industry category (consider healthcare versus automobile retailing), competitive, mimetic, or normative pressures to adopt more equitable employment practices will likely lead to greater representation of women at higher levels across firms (Dobbin, 2009; Stainback et al., 2010). Thus the representation of women in executive ranks within an industry is a reflection of greater scrutiny from regulatory bodies and employee groups; it is also a reflection of the diffusion of egalitarian employment practices in that industry. These attributes of the context should also facilitate equitable employment outcomes for men and women (Dobbin, 2009). Conversely, industries in which women have not been able to break into the upper echelons reflect contexts in which institutional pressures to adopt equitable practices have been less salient. In these settings, biased performance reviews or bonus and promotion decisions favoring men over women are more likely. Based on these arguments, we propose that:

Hypothesis 3a. The proportion of women at executive and senior managerial levels within a specific industry reduces sex differences in performance evaluations such that women receive equal to, or higher, performance evaluations than men in industries with a higher proportion of women at executive or managerial levels.

Hypothesis 3b. The proportion of women at executive and senior managerial levels within a specific industry reduces sex differences in organizational rewards such that women receive equal to, or higher, rewards than men in industries with a higher proportion of women at executive or managerial levels.

Job-Level Effects on Sex Differences in Performance Evaluations and Rewards

The job context encompasses a number of mechanisms by which male–female differences in employment outcomes manifest in organizations. For instance, the structure of jobs drives how incumbents

relate to each other and compare their employment circumstances. The structure of jobs also drives the definition and weighting of criteria for evaluating and rewarding incumbents (Lansberg, 1989). We focus on job complexity—i.e., the extent to which jobs involve a high level of problem-solving skills, discretion, and autonomy—as an important facet of the job context (Auster, 1989).

Research suggests that, in complex job settings, the cognitive demands placed on supervisors when evaluating performance and rewarding employees are fairly high; in these contexts, cognitive biases and stereotyping are more likely to play a role in supervisory performance evaluations and reward allocation (Auster, 1989; Fiske et al., 1991). In addition, discrimination and bias in evaluation is more likely to occur when information about evaluating a job cannot be easily collected and documented (Heilman, 1994; Nieva & Gutek, 1980). Some researchers note that employers may be more likely to behave stereotypically when there is limited, deniable, and ambiguous information regarding an individual's performance (Gerdes & Garber, 1983; Mobley, 1982). Complex jobs represent such settings. For instance, Mobley (1982) studied the performance appraisals of nonmanagement employees in a large supply organization, and found that evaluations of females in nonprofessional and nonmanagerial jobs are less likely to be affected by systematic gender bias owing to the lower complexity of nonmanagerial jobs.

Although the ambiguity or subjectivity of job performance criteria and the related propensity to stereotype against women is one important mechanism by which job complexity might enhance sex differences in performance evaluations, quantitative and narrative reviews suggest that subjective performance evaluation criteria alone do not always disadvantage women (Bartol, 1999; Roth, Purvis, & Bobko, 2012). We therefore posit that the ambiguity or subjectivity of the evaluation criteria may not be the only mechanism by which sex differences in performance evaluations or rewards emerge in complex jobs; rather, the complexity of the job context may shape subtler forms of bias that lead to systematic sex differences in these outcomes. For instance, complex jobs often involve more variable and demanding work schedules, as well as expectations that employees will put in “face time” after working hours. It is possible that although women are not necessarily disadvantaged in terms of performing the core tasks of a job, other intangible requirements of complex jobs might not work in their favor (Cleveland et al., 2005; Correll, Kelly, O'Connor, & Williams, 2014). Moreover, complex jobs are also

jobs with higher status and span of control in organizations (e.g., managerial and executive positions). Following the role congruence perspective outlined earlier (e.g., Eagly et al., 1992), it is also likely that women in these jobs will be considered atypical and face reward penalties.

Finally, job complexity is also associated with greater role autonomy, discretion, and idiosyncratic bargaining by incumbents, leading to an additional mechanism by which job complexity drives gender inequality: lack of opportunities for incumbents to compare objective employment conditions (Baron & Pfeffer, 1994). In complex jobs, because specific roles and terms and conditions of employment are less generalizable across incumbents, individuals are less likely to detect or report unequal treatment (Lansberg, 1989). Together, these perspectives suggest that job complexity is an important embedding context governing reward allocation and evaluative decisions in organizations. Thus we propose that:

Hypothesis 4a. Job complexity enhances sex differences in performance evaluations such that men receive higher performance evaluations than women in more complex jobs.

Hypothesis 4b. Job complexity enhances sex differences in organizational rewards such that men receive higher organizational rewards than women in more complex jobs.

Performance Differences as Explanations for Reward Differences between Men and Women

Although we have argued that sex differences in rewards and performance evaluations are primarily a function of “where” and “what” work is being performed, we also acknowledge that gender-based differences in performance evaluations may explain sex differences in organizational rewards. In field settings, studies that have accounted for performance differences in explaining reward differences are rare. Past research has shown that women receive fewer rewards in organizations even after accounting for attributes that might underlie performance differences, such as educational level and labor market experience (Lyness & Thompson, 1997; Morrison & Von Glinow, 1990; Stroh, Brett, & Reilly, 1992).

More recently, some researchers have directly accounted for performance differences as explanations for reward differences among men and women. In the context of sales employees working for a single firm, Joshi and colleagues (2006) found that objective

sales performance (i.e., sales targets achieved) partially explained sex differences in annual salary after accounting for the sex composition of the work unit. In a more recent longitudinal analysis of personnel practices in an organization in the service industry, Castilla (2012) reported that sex and race differences in salary increases persisted after controlling for performance evaluations received by employees (see also Elvira & Town, 2001). In fact, researchers have found that reward systems that are tied to performance evaluations tend to exacerbate (rather than explain) inequity based on demographic attributes (Castilla, 2012; Castilla & Benard, 2010). Contrary to these findings, however, across employment settings, Blau and Devaro (2007) found that, after controlling for performance differences, there were no significant difference in wage growth between men and women. In order to empirically examine whether these performance differences explain gender inequality in organizations across multiple job contexts and to resolve discrepancies in prior findings, we posit that:

Hypothesis 5. Differences in organizational rewards can be fully accounted for by differences in performance evaluations between men and women.

Bundles of Contextual Factors Associated with Sex Differences in Performance Evaluations and Rewards

So far, based on extant research, we have proposed the independent effect of occupation-, industry-, and job-level variables on performance and reward differences between men and women. Based on past research, it is not clear whether two or more of these contextual attributes might operate together to explain sex differences in work-related outcomes. In order to develop a unified framework, it may also be worthwhile to consider whether these various attributes jointly influence sex differences in employment outcomes. For example, do gender-balanced occupations *and* industries with a high proportion of female executive jointly enable women to close the gender gap in organizational rewards? Are men more likely to outperform women while performing complex jobs in prestigious occupational contexts? In order to examine whether bundles of occupation-, industry-, and job-level contexts combine or complement each other to drive sex differences in performance and rewards, we also conducted an exploratory qualitative comparative analysis (QCA), as described below.

METHODS

Literature Search

We used computerized database searches of *PsycINFO*, *EBSCO*, and *ProQuest* to generate a pool of potential articles. To identify all of the articles that investigated gender and our outcomes of interest, we used combinations of the following search terms: “gender” (“sex,” “male/female”), and “task,” “job,” “in-role performance,” or reward outcomes such as “pay,” “promotion,” “salary,” “wages,” “rewards,” or “bonus.” We also manually searched major management journals in the field (Gomez-Mejia & Balkin, 1992), as well as disciplinary journals that focus on organizational research, to supplement the electronic search (e.g., *American Journal of Sociology*, *Industrial and Labor Relations Review*). To find unpublished sources such as dissertations and convention presentations, we searched Dissertation Abstracts International, and the programs of the annual meetings of the Academy of Management and the Society for Industrial and Organizational Psychology. We also contacted researchers in related areas to obtain current and unpublished studies that might fit our criteria for inclusion.

We used several inclusion criteria to select studies in the meta-analysis. First, we included studies that reported an *individual-level* effect of gender on performance evaluations or rewards. To calculate the standardized mean difference (d) between male and female groups, a study had to report the mean, standard deviation, and sample size for each gender group. If no such information was available, an appropriate statistic such as zero-order correlation, t , or F had to be provided to allow the computation of the standardized mean difference using formulas detailed by Lipsey and Wilson (2001). We also included studies that considered gender as a control variable as long as they reported relevant information with respect to performance and rewards (e.g., correlations) and transformed into the standardized mean difference. Because our study examined the moderating effects of occupation-, industry-, and job-level context on gender inequality, we focused primarily on studies that were conducted in field settings in which the effects of these variables were likely to manifest in sex differences in performance and rewards. We did not include in our dataset studies that relied on student samples, used experimental designs, or involved tasks in artificial environments (e.g., simulations).

Using all of the processes described above, the final sample set of our analyses yielded a total of 190 effect sizes and 474,732 individuals from 142 studies (73 for performance evaluations and 69 for organizational

rewards) conducted between 1985 and 2013.¹ This time period covers almost 30 years of research on the topic of gender in the management domain, and also represents the era following the civil rights legislation and the feminist movement, during which time women have made substantial gains in education, but have continued to face challenges with respect to gender parity in the workplace (Blau & Kahn, 2007; Reskin, 2003). From a practical standpoint, the reporting of correlations and other descriptive statistics was sparse in the research prior to the 1980s.

Coding and Variables

We thoroughly examined and coded all studies included in the sample based on a coding protocol (Lipsey & Wilson, 2001). We collected relevant information regarding effect size (d), key antecedents (occupation, industry, job level), and control variables from sample studies. We also collected data from secondary sources for our moderating variables (see below for a detailed explanation). Given that we used objective data for all of the variables included in the analyses, subjectivity or intercoder unreliability was not a concern; however, we also cross-checked our coding with all of the coders several times throughout the coding process.

Performance evaluations. Measures of *performance evaluations* indicated a rater's (mostly a supervisor's) evaluation of an employee's overall effectiveness and performance in fulfilling his or her everyday roles and responsibilities (e.g., Wayne & Liden, 1995; Williams & Anderson, 1991). We focused on job, task, or in-role performance (Borman & Motowidlo, 1997), and excluded studies that measured other dimensions such as contextual performance, citizenship behavior, or creative performance. When multiple measures of performance evaluations were available from a single study, we calculated a composite effect size by averaging all effect size information. We also coded the reliabilities of the measurement instruments whenever available. When the reliabilities were not reported, we used the average reliability of the same variable from all of the other studies in our dataset. The average reliability of the performance evaluation measure was .89.

Organizational rewards. Consistent with the conceptualization of tangible or actual rewards in the previous meta-analyses on this topic (Ng, Eby, Sorensen, & Feldman, 2005; Podsakoff, Whiting,

¹ A brief summary of studies included in the analysis is available from the first author on request.

Podsakoff, & Blume, 2009), in our study we defined *organizational rewards* as extrinsic forms of reinforcement offered by an organization (Byron & Khazanchi, 2012). Specific measures included salary level and salary increases, bonuses and incentive payments, and number of promotions. All of the studies in our dataset provided objective organizational rewards measures.

Proportion of men in an occupation. To obtain data regarding the *proportion of men in an occupation*, we referred to the U.S. Bureau of Labor Statistics (BLS) Current Population Survey. We first coded relevant information (occupations involved in the samples) based on sample descriptions of primary studies; we then obtained the occupational sex composition data from the BLS dataset and assigned this information to the occupations included in the studies as close to the year in which the primary study was conducted. For studies that included multiple occupations, we found sex composition data for each occupation category from the BLS and then calculated a composite value by averaging information on all participating occupations. However, we excluded from the analysis any studies that involved multiple occupations with varied sex composition (e.g., accounting and engineering in the same sample), involved non-U.S. samples, or were unclear about the samples ($n = 38$). Same coding rules were applied to coding other key variables.² Among the studies

included in our analysis, the percentage of males in various occupations ranged from 12% (i.e., bank tellers) to 88% (i.e., automotive dealers) and 97% (i.e., firefighters), with an average of 61%.

Occupational prestige. *Occupational prestige* refers to the consensual rating of an occupational category with respect to its worthiness, based on several socioeconomic factors such as income and the education levels of incumbents (Hauser & Warren, 1997). To code and obtain information for occupational prestige, we relied on the Nakao-Treas Socioeconomic Index (Nakao & Treas, 1994), which is based on a nationwide survey undertaken by the National Opinion Research Center (NORC). It provides prestige scores of a wide range of more than 500 occupation titles ranked on a 100-point scale. Specifically, it uses socioeconomic indices such as education and income levels that are standardized for full-time incumbents' age distributions. The prestige scores are then calculated based on the age-weighted average of education and income levels for each occupation. We first coded occupation titles based on the sample descriptions in primary studies and then assigned the corresponding occupational prestige score from the index. The mean of the Nakao-Treas prestige scores in our sample was 53.2—equivalent to a middle-class occupational prestige level. Examples of less prestigious occupations in our sample were seasonal sales clerks and warehouse workers, with scores of 27.7 and 22.1, respectively. Examples of the most prestigious occupations in our sample were executives and primary care physicians, with scores of 64.2 and 71.79, respectively.

Proportion of women at executive or senior manager levels within industry. To code the *proportion of women at executive or senior manager levels within industry*, an industry-level context variable, we consulted the upper level (executives and senior managers) female representation data for each industry available in the U. S. Equal Employment Opportunity Commission (EEOC) survey, *Job Patterns for Minorities and Women in Private Industry Survey*. Drawing on detailed information on the EEOC survey, which was based on the North American Industry Classification System (NAICS) two-digit industry code, we obtained data about the female representation in upper level positions for the industries identified and included in the sample. Once again, we obtained data as close to the timing of the primary studies as possible and used only U.S.-based samples. Examples of industries with fewer high-ranked women were utility (13%)

² In order not to oversimplify the coding of job, industry, and occupational variables, we took the following steps during the coding process. For primary studies that included samples within a single work context—e.g., primary care physicians (Hekman, Aquino, Owens, Mitchell, Schilpzand, & Leavitt, 2010), public school teachers (Ferris, Bergin, & Wayne, 1988), executive managers (Graves, Ruderman, Ohlott, & Weber, 2012), bank tellers (Schaubroeck & Lam, 2002), automotive dealers (Liu, Loi, & Lam, 2011), etc.—we were able to code the key variables directly. These studies constituted more than 50% of our data. Primary studies that relied on multiple organizations or work contexts often included occupations and jobs that were very similar, e.g., scientists and engineers (Keller, 2012), system analysts and programmers (Igbaria & Baroudi, 1995), etc. In these cases, we used average occupational gender composition, occupational prestige, or job complexity scores. However, we excluded from our regression analysis studies that provided relatively limited information about the sample or which involved multiple diverse work contexts—e.g., a mixed sample consisting of employees working in areas of finance, sewage, libraries, and childcare (Jepsen & Rodwell, 2009)—or non-U.S. samples.

and manufacturing (18%); on the other hand, the healthcare industry employed about 54% of upper level female managers, which was the highest female representation at executive levels in the sample.

Job complexity. We obtained *job complexity* scores using scores available at the *Occupational Network (O*Net)* database—a source that contains information on standardized and job-specific descriptions of the work activities, content, characteristics, and skills required for 975 job titles. Based on operationalizations in past research (Auster, 1989; Campbell, 1988), job complexity was assessed with the *O*Net* skill variable labeled “complex problem solving,” which indicates the abilities needed to solve novel, ill-defined problems in complex, real-world settings, and is typically higher for jobs that are characterized by greater discretion and autonomy. The *O*Net* values for complex problem-solving skill ranged from 0 to 100, with higher scores indicating the skill’s greater importance to the job. Our sample includes relatively less challenging and more routine jobs such as clerical workers (38) and bank tellers (47), as well as highly complex jobs such as senior scientists or engineers (85) and top executives (88).

Control variables. We also coded several control variables (study and sample characteristics) to enhance the robustness of the analyses. First, we used a dummy variable capturing whether the effect sizes were obtained from studies with longitudinal (coded as “1”) or cross-sectional (coded as “0”) designs. Publication status was also dummy coded (“1” if published). To take into account potential temporal effects, we coded the publication year of each primary study from which effect sizes were derived (Carney, Gedajlovic, Heugens, Van Essen, & Van Oosterhout, 2011). We coded average sample tenure from each study (the average number of years that individuals spent in the organization) because tenure can serve as a proxy for the human capital of employees included in the sample (e.g., Dipboye & Colella, 2005). We also considered rating source (i.e., supervisory versus peer ratings) as an additional control; however, it was not included in a final model because of a lack of studies using peer ratings in our data.

Meta-Analytic Procedures

We employed several types of analytical approach to the meta-analysis. First, to examine sex differences in performance and rewards, we calculated effect

size d , the standardized mean difference between female and male groups for continuous measures of performance evaluations and organizational rewards, using procedures suggested by Hedges and Olkin (1985). Measurement unreliability was corrected for subjective performance ratings (Lipsey & Wilson, 2001). We calculated confidence intervals of 95% to judge the statistical significance of the effect sizes (Whitener, 1990). To test the homogeneous distribution of the effect sizes (detecting possibilities of potential moderators), we also calculated the Q statistics (Hedges & Olkin, 1985).

Testing for potential publication bias. To address any potential publication bias issues, we generated contour-enhanced funnel plots (Kepes, Banks, McDaniel, & Whetzel, 2012). These graphs are scatterplots in which the effect sizes estimated from primary studies are displayed on the X axis and the inverse of a sample’s standard error (i.e., precision) along the Y axis (Kepes et al., 2012). Since the precision depicted on the Y axis increases as sample size increases, in the absence of publication bias, studies with larger (i.e., more precise) sample size are scattered toward the top, whereas studies with small sample size (thus less precise) are scattered widely toward the bottom of the funnel plot. If sampling error is not the only reason for variance in the sample distribution and studies with insignificant results are omitted (i.e., the presence of publication bias), asymmetrical funnel plots may result (Kepes et al., 2012); if the funnel plot is symmetrical, we may conclude that there is no publication bias.

Meta-regression analysis. To test our main hypotheses (Hypotheses 1–4), we conducted random effects meta-analytic regression analysis, which comprises a series of weighted least squares (WLS) regressions investigating the relationship between key independent variables and the effect estimate as the outcome variable (Lipsey & Wilson, 2001). In random effects meta-regressions, effect sizes are weighted by the differences in precision (inverse variance weights) (Hedges & Olkin, 1985). This approach also allows us to take into account residual heterogeneity not modeled by the independent variables, but arising out of unmeasured between-study differences (Lipsey & Wilson, 2001). We tested the effects of each of the occupation-, industry-, and job-level variables with control variables in separate models. Although meta-analytic regressions allow us to test the effect of multiple factors (i.e., all independent variables) at the same time in a full regression equation, this approach is

often not feasible because of insufficient studies across all variables (Thompson & Higgins, 2002).

Monte Carlo simulations for testing robustness of estimates. To check the robustness of our regression findings, we further conducted a permutation test based on Monte Carlo simulations (Higgins & Thompson, 2004). In meta-regressions, owing to relatively small sample sizes (especially when multiple covariates are considered), there is an increased chance of at least one false positive finding (i.e., of “Type I” error) (Harbord & Higgins, 2008). In permutation tests, the covariates are randomly allocated to the outcomes over multiple runs (1,000 runs in our analyses) and a *t*-statistic is calculated for each run. The true *p* value for the relationship between a given covariate and the outcome is computed by counting the number of times that these *t*-statistics are greater than or equal to the observed *t*-statistic. Thus the adjusted *p* value from a permutation test can be interpreted as the degree to which one may be surprised about the observed effect of a variable if all other covariates were also included in the meta-regression (Higgins & Thompson, 2004). A permutation test approach is highly recommended to assess statistical significance in meta-regressions (see Higgins & Thompson, 2004 for details).

Mediation test. To test whether sex differences in rewards can be explained by performance evaluation differences (Hypothesis 5), we conducted a semipartial correlation test based on the meta-analytic estimates calculated about $X \rightarrow Z$, $X \rightarrow Y$, and $Z \rightarrow Y$ (where *X* is gender, *Z* is performance evaluations, and *Y* is organizational rewards). (For a similar approach, see Gajendran & Harrison, 2007.) We first transformed our standardized mean difference (*d*) to correlation (*r*) using the equation provided by Lipsey and Wilson (2001). After controlling for the effect of performance evaluations, we calculated a semipartial correlation between gender and rewards using the harmonic mean as the sample size for the meta-analytic correlation (Viswesvaran & Ones, 1995). We used semipartial correlations instead of partial correlations because the former provides a correlation such that performance (*Z*) is partialled out of gender (*X*) alone, whereas a partial correlation coefficient has performance (*Z*) partialled out of both rewards (*Y*) and gender (*X*) (Lewis-Beck, Bryman, & Liao, 2004). Given that we were interested in the change in $X \rightarrow Y$ after partialling out $X \rightarrow Z$, not $Z \rightarrow Y$, the semipartial correlation could give us a precise statistic for testing mediation by not holding performance constant for rewards.

Exploratory qualitative comparative analysis to identify bundles of contextual factors. We applied a crisp set version of the QCA (Ragin, 2000; 2008) that has received growing attention in management research and which is ideal for identifying how multiple independent variables can jointly influence a phenomenon of interest (see Fiss, 2011; Greckhamer, Misangyi, Elms, & Lacey, 2008). Unlike other statistical procedures that rely on significance testing or aim at explaining variance in an outcome of interest, QCA identifies which specific set of causal attributes is common across all cases of an outcome. For example, the analysis could identify the combinations of occupational, industry, or job attributes that occur across cases of higher levels of sex differences in rewards or performance (Greckhamer et al., 2008). In our study, a “case” is an effect size (*d*) derived from primary studies. We applied QCA in this study as an exploratory approach for identifying whether multiple attributes (i.e., occupation-, industry-, and job-level variables) are associated with cases of high levels of sex differences in rewards or performance in the data. Qualitative comparative analysis relies on the principles of Boolean algebra, and utilizes binary data and combinatorial logic to identify the causally relevant conditions that explain a focal phenomenon (Ragin, 2008).

A crisp set QCA analysis is inherently dichotomous and evaluates each case in the data based on either membership or nonmembership of a set (see Greckhamer et al., 2008, for details). The cutoffs for dichotomizing attributes were based on theoretical considerations (Ragin, 2000). Conceptually, a *d* greater than zero implies a set of cases in which men receive higher evaluations or are rewarded more than women, while a *d* lower than zero represents cases in which women receive higher evaluations or are rewarded more than men. Although calibrating outcomes based on theory alone is ideal, practical considerations, such as the distribution of data, can also guide these decisions (Greckhamer et al., 2008). For performance differences, our data included a full range of effect sizes, with *ds* ranging from positive, near zero, to negative, with a median close to zero (i.e., *ds* ranged from -2.16 to 2.27 , with a median of $-.01$). Based on conventions followed in past research, as well as theoretical considerations, we defined a set of cases in which *ds* were positive as cases in which men received higher evaluations than women and other cases as nonmembers of this set. We also defined a set of cases in which *ds* were lower than the median value as

a set of cases in which women received higher evaluations than men and other cases as nonmembers of this set. For reward differences, d s were positively skewed, ranging from $-.37$ to 3.07 , with a median of $.65$. We used the median as a cut-off point to represent a set of cases in which men were rewarded higher than women and other cases as nonmembers of this set.

Each occupational, industry, and job attribute was also similarly calibrated, with cases above the median value coded as “1,” meaning the presence of a specific condition, or “0,” meaning nonmembers of this group. Based on median values, each of the contextual attributes was defined as follows: Occupations with more than 40% women were categorized as “balanced occupations,” occupations with prestige scores of more than 60 were classified as “prestigious occupations,” industries with more than 30% female executives or senior managers were categorized as “high proportion of women in executive/senior manager levels,” and jobs with complexity scores over 70 were classified as “complex jobs.”

Once the data were coded into crisp sets, the second phase of the QCA approach involved examining the sufficiency of the four attributes (occupational demographic composition, occupational prestige, percentage of female executives in the industry, and job complexity) for explaining the four outcomes: Men evaluated higher than women; women evaluated same as, or higher, than men; men rewarded higher than women; women rewarded same as, or higher, than men. Based on the truth table algorithm provided by the QCA software, configurations for each of the outcomes were specified. In the current analyses, configurations that yielded at least three cases with consistency of more than 1.00 were included in the final analysis. These cut-off values are also based on conventions followed in past QCA research. Our decision to include only configurations with at least three cases allowed us to include 80% of the cases in the analyses (see Misangyi & Acharya, 2014; Ragin & Rihoux, 2009). We used the most stringent cut-off of 1.00 for the consistency of the cases represented by that configuration (Ragin, 2000). Consistency is a measure of sufficiency of a particular configuration of attribute, and it indicates how closely a configuration is associated with an outcome. A consistency of 1.00 indicates that every time a particular combination of attributes is present, the outcome is also present. In other words, a configuration was included in the final results only when

100% of cases displaying a specific configuration of occupation-, industry-, or job-level attributes also displayed the outcome of interest (Ragin, 2008).

RESULTS

Sex Differences in Performance Evaluations and Organizational Rewards

Using the meta-analytic techniques described above, we first examined the magnitude of sex differences in performance evaluations and rewards (see Table 1). For performance evaluations, the mean effect size corrected for unreliability was positive (i.e., men outperformed women), but the confidence interval included zero ($d = .04$, $k = 93$, 95% confidence interval, or CI = $-.12$ – $.20$), indicating that there were no significant sex differences in performance evaluations. For organizational rewards, results indicated that men received significantly higher pay and promotions than women ($d = .56$, $k = 97$, 95% CI = $.45$ – $.65$).³ The mean sex difference in rewards was 14 times larger than the mean sex difference in performance evaluations. Table 1 also shows that there is considerable heterogeneity among the effect sizes, as indicated by the Q statistic. The Q values for both performance evaluation and reward outcome measures were highly significant ($p < .01$), indicating that the effect sizes vary across the studies and that potential moderators play a role.

To check potential publication bias in our results, we also plotted contour-enhanced funnel plots for both performance and reward outcomes. In the plot, the white area is where statistically insignificant effect sizes are located. The darkest and thinnest area is where marginally significant effect sizes lie ($.05 < p < .10$). The thin and dark gray shaded, and the large light gray shaded, areas are where statistically significant effect sizes are found (Kepes et al., 2012). For performance evaluations, the plot is symmetrical (see Figure 1a); also the dots, which

³ Since we grouped together various types of reward (i.e., pay, bonus, and promotion) into a broad category of organizational rewards, we also undertook additional subgroup analyses to see whether there were any differences based on the type of reward. The overall patterns were consistent, and we found that sex differences were larger for pay ($d = .67$, $k = 67$, 95% CI = $.55$ – $.75$) than for promotions ($d = .24$, $k = 29$, 95% CI = $.08$ – $.39$). Owing to the lack of samples, we were not able to conduct the separate analyses for each subgroup using more specified types of reward.

TABLE 1
Sex Differences in Performance Evaluations and Organizational Rewards

Outcomes	<i>k</i>	<i>n</i>	Effect size <i>d</i>	95% CI	<i>Q</i>
Performance evaluations	93	95,882	0.04	−0.12, 0.20	53,545.10*
Organizational rewards	97	378,850	0.56	0.45, 0.65	207,187.50*

Note: *k* = no. of effect sizes; *n* = total no. of individuals counted by effect sizes; effect size *d* = sample size weighted mean effect size corrected for unreliability; 95% CI = 95% of confidence interval of *d*; *Q* = effect size heterogeneity statistic indicating the possibility of moderators

**p* < .01

represent studies included in the analysis, are evenly distributed across the shadowed areas (i.e., regions of statistical significance), as well as the white area (i.e., the area of nonsignificance). For rewards (see Figure 1b), the plot is less symmetrical, indicating the possibility of the skewedness of our samples. However, a closer look at the plot reveals that the dots are allocated across regions of statistical significance, as well as regions of nonsignificance (the white area). Moreover, potentially “missing” studies are more likely to be located on the bottom left-hand side of the plot, which is the region representing statistical significance, but lower sample sizes or higher standard errors. (Note that there are almost no dots in that area.) Thus, for both performance and reward outcomes, we did not find strong publication bias issues in the sample.

Results of Meta-Analytic Regressions

Table 2 summarizes results from a series of meta-analytic regressions for performance evaluations (Models 1–3) and reward differences (Models 4–6).⁴ Overall, we found general support for the hypotheses. Hypotheses 1a and 1b predicted that sex differences in performance evaluations and rewards would increase as the proportion of men in an occupational

category increased. In support for the hypotheses, we found significant positive relationships between the proportion of men in an occupation and performance differences (Model 1: $b = 1.37$, $p < .01$), as well as reward differences (Model 4: $b = 2.08$, $p < .01$), between men and women.

Our second set of hypotheses (Hypotheses 2a and 2b) predicted that sex differences in performance evaluations and rewards would increase as the prestige level of the occupational context increases. Models 1 and 4 also present the regression results testing these hypotheses. Contrary to our expectations, we did not find a significant relationship between occupational prestige and sex differences in performance evaluations (Model 2: $b = .03$, $p > .10$); however, sex differences in organizational rewards increased with occupational prestige (Model 4: $b = .06$, $p < .05$), rendering support for Hypothesis 2b. Both regression models were statistically significant (Model 1: $F = 5.68$, $p < .05$; Model 4: $F = 6.15$, $p < .01$), and the adjusted R^2 also indicated that more than 20% of between-study variance can be explained by occupational sex composition and prestige, along with other control variables.

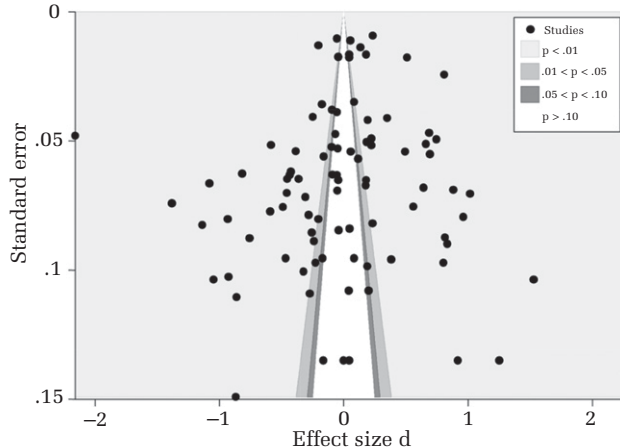
Hypotheses 3a and 3b predicted that sex differences in performance and rewards would be smaller in industries with a higher representation of women at the executive and senior managerial levels. Models 2 and 5 in Table 2 represent these findings. Providing strong support for our hypotheses, results revealed that there were significant negative relationships between the representation of women at higher levels within industry and performance evaluation differences (Model 2: $b = -2.80$, $p < .01$), as well as reward differences (Model 5: $b = -3.34$, $p < .01$). Both regression models were significant (Model 2: $p < .01$, adjusted $R^2 = .18$; Model 5: $p < .01$, adjusted $R^2 = .23$).

As posited in Hypotheses 4a and 4b, we found that job complexity had a positive effect on sex differences in performance evaluations and rewards. Findings indicated that men received higher performance evaluations as job complexity increased (Model

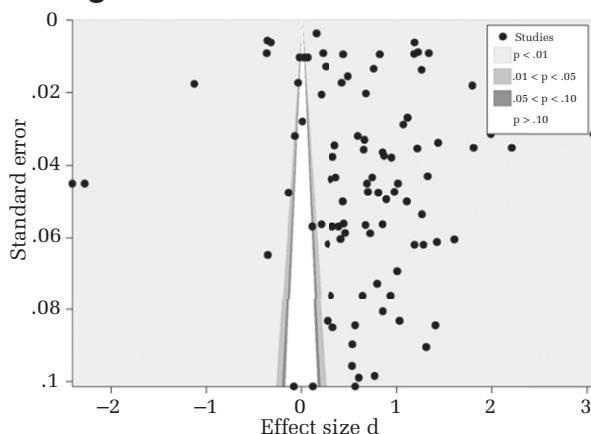
⁴ We present results for occupation-, industry-, and job-level variables, along with control variables in this section. As a result of the listwise deletion of missing cases, including all variables in a single meta-regression was not feasible. However, we undertook additional analyses that included various combinations of occupation- and industry-level variables simultaneously to examine whether the overall pattern of results remained the same and explained additional overall variance in the effect sizes. For both performance evaluation and reward outcomes, we found consistent patterns of results in our multiple regressions. Variance inflation factors for these analyses were well below the cut-off value of 10. We also present results of Monte Carlo simulations that address any concerns of Type I errors.

FIGURE 1
Funnel Plot for Studies Testing Sex Differences in
Performance Evaluations and Organizational
Rewards

A. Performance Evaluations



B. Organizational Rewards



3: $b = .06$, $p < .01$). The relationship between job complexity and reward differences between men and women also revealed a similar pattern (Model 6: $b = .05$, $p < .01$). Both regression models were also statistically significant (Model 3: $p < .01$, adjusted $R^2 = .17$; Model 6: $p < .01$, adjusted $R^2 = .23$).

Results of Monte Carlo simulations. Based on Monte Carlo simulations, we conducted permutation tests to ascertain the robustness of these results. After 1,000 random runs based on Monte Carlo simulations, adjusted p values for significant effects remained significant ($p < .05$ for industry upper level female composition with performance evaluation differences; $p < .01$ for all others) after accounting for multiplicity (multiple testing), indicating the robustness or low probability of Type I error in our findings. For example, the adjusted

p value of .008 for occupational prestige with reward differences (more conservative than $p = .003$, as per Model 6) indicates a lower than 1% probability of false positive error; the adjusted p value for industry upper-level female composition with performance differences increases from .006 to .024 after 1,000 random runs, but it still remains significant ($p < .05$).

Mediation Test: Semipartial Correlation

To test whether sex differences in performance evaluations explain sex differences in organizational rewards (Hypothesis 5), we conducted a semipartial correlation test. Table 3 reports a semipartial correlation between sex and organizational rewards after controlling for sex differences in performance evaluations. Results indicated that performance evaluations did not appear to mediate the effects of sex on organizational rewards when comparing the semipartial correlation ($r = .27$, $p < .01$) with the original correlation ($r = .27$, $p < .01$). Sex differences in rewards remained significant after partialling out the effects of performance evaluations.

Results of Exploratory QCA Analyses

Table 4 presents the findings of a crisp set QCA analysis.⁵ The configurations are based on conventions developed in recent research (see Fiss, 2011; Greckhamer et al., 2008). Central contextual attributes are represented by • if present and ⊗ if absent. Peripheral attributes are represented by • if present and ⊗ if absent. Central conditions do not require

⁵ The following are illustrations of each of the four configurations from our data: Panel 1—Automotive dealers in the automobile retail industry engaged in direct customer sales (both the occupation and industry executive ranks are male-dominated and the occupation is rated as less prestigious); Panel 2:1—Bank tellers working in the retail banking industry (jobs are routine and the occupation is balanced, but relatively few women are represented at executive levels in the industry); Panel 2:2—Security employees in the retail services industry (e.g., malls), which task mostly entails assisting mall customers (the occupation is male-dominated, but there is a relatively high proportion of women at the executive levels in this industry); Panel 3—Engineering consultants working in the utility industry (the occupation is prestigious and male-dominated, senior executive ranks in the utility industry are also male-dominated, and tasks are dynamic and complex in nature).

TABLE 2
Results of Meta-Analytic Regressions

	Performance evaluations			Rewards		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Constant	-0.61 (0.36) [†]	0.82 (0.43) [†]	-1.29 (0.49)*	-0.69 (0.47)	1.45 (0.38)**	-2.50 (0.76)**
<i>Controls</i>						
Study design (1 = longitudinal)	-0.06 (0.24)	-0.14 (0.26)	-0.06 (0.25)	-0.04 (0.20)	-0.03 (0.18)	-0.07 (0.19)
Publication year	-0.02 (0.01)	-0.00 (0.01)	-0.01 (0.01)	-0.03 (0.01) [†]	-0.00 (0.01)	-0.01 (0.01)
Average sample tenure	0.02 (0.02)	0.03 (0.03)	0.02 (0.02)	-0.03 (0.03)	-0.01 (0.02)	-0.02 (0.03)
<i>Predictors</i>						
Proportion of men in occupation	1.37 (0.40)**			2.08(0.65)**		
Occupational prestige	0.03 (0.01)			0.06 (0.01)*		
Proportion of women in executive or senior manager levels within industry		-2.80 (0.74)**			-3.34 (0.85)**	
Job complexity			0.06 (0.01)**			0.05 (0.01)**
<i>k</i>	58	54	61	71	44	74
Adjusted <i>R</i> ²	0.21	0.18	0.17	0.22	0.23	0.23
Model <i>F</i>	5.68*	3.79**	3.90**	6.15**	4.13**	6.49**

Note: Unstandardized regression coefficients are presented; numbers in parentheses are standard errors; *k* = total no. of effect sizes involved in the analysis; the publication status (1 = published) dummy was dropped from the analysis owing to multicollinearity

[†] *p* < .10

* *p* < .05

** *p* < .01

counterfactual analyses and can be fully determined by the cases in the data. Peripheral conditions, on the other hand, require counterfactual analyses, and their inclusion in the solution is based on theoretically derived assumptions. Counterfactual analyses require all possible configurations (in our study, 2^k where *k* = four attributes) to be populated with cases, although it is often unlikely that all logically possible configurations can be populated in real organizational contexts (Ragin, 2008).

Further, each solution may consist of multiple configurations of attributes, and “coverage” of each configuration shows the proportion of consistent cases that display that particular configuration of attributes. Unique coverage is the proportion of cases that display only that particular configuration. In addition, the analysis also provides the overall consistency and coverage for all of the configurations in a particular solution. Table 4 shows three panels displaying solutions for each of the outcomes of interest. We excluded one outcome from the analyses—women rewarded the same, or higher, than men—because there were insufficient cases in the data. As shown in Panel 1, one configuration was sufficient for men receiving higher performance evaluations than women. Men received higher evaluations than women in male-dominated occupations (i.e., absence of gender-balanced

occupation) that were less prestigious (e.g., automotive sales, assembly line manufacturing) and had few women at the executive levels in the industry regardless of whether the job was complex or routine.

Panel 2 shows that two configurations were sufficient for women receiving higher performance evaluations than men. Across both configurations, note that “job complexity” is clearly absent. Further, gender balance in the occupation and higher proportion of women at senior executive or senior manager levels appear to be substitutes for one another: When one is present, the other is absent, and vice versa. Therefore it appears that women employed in routine jobs are likely to receive higher evaluations than men when either the occupation in which they are working is balanced or the industry as a whole has a higher representation of women at executive levels. These combinations apply regardless of occupational prestige.

Panel 3 shows that one configuration was sufficient for men receiving higher rewards than women. Men received higher rewards than women in prestigious occupations in which jobs were complex and women were underrepresented in senior positions at the industry level. Whether the occupation was balanced or not was irrelevant. Overall, across both performance evaluation and reward outcomes,

TABLE 3
Mediation Test: A Semipartial Correlation

Outcome	Estimated <i>r</i>	Semipartial correlation after controlling for differences in performance evaluations
Organizational rewards	0.27** (378,850)	0.27** (60,259)

Note: Estimated *r* indicates a correlation corrected for unreliability; significance of a semipartial correlation was evaluated against the harmonic mean of sample sizes (in parentheses)

***p* < .01

it appears that higher representation of women at executive levels in the industry (presence or absence) was a common condition. Gender balance at the occupational level and job complexity emerged as common conditions in three out of the four configurations across all outcomes.

DISCUSSION

Nearly five decades after the passage of the Civil Rights Act of 1964, academics and policymakers are still asking the question: Can women close the gap in performance and rewards in the workplace? The answer to this question based on our research is: Women often close the gap in performance, but not the gap in rewards. Across occupations ranging from bank tellers to accountants, industries ranging from IT to healthcare, and jobs ranging from mundane to challenging, our results show that sex differences in organizational rewards were almost 14 times larger than sex differences in performance evaluations. Moreover, performance differences did not explain reward differences between men and women. The percentage of men in an occupation and the complexity of jobs performed enhanced the male–female gap in performance and rewards. In highly prestigious occupations, women performed equally, but were rewarded significantly lower than men. Only in industries with a higher proportion of female executives did women reverse the gap in performance evaluations *and* in rewards.

Theoretical Implications: Identifying Pathways to Gender (In)Equity in Organizations

Our meta-analysis provides broad insights into the extent to which reward and performance differences vary based on various macro-level attributes of the work context, and we suggest below how future studies might further unpack pathways through which these contextual factors shape the

allocation of rewards and performance evaluations among men and women.

A noteworthy macro–micro linkage in our findings is represented by the significant effects of fairly distal contextual attributes—occupational sex composition and prestige—on differences in performance evaluations and rewards among men and women. In male-dominated occupations, it is possible that the masculine cultural norms lay the groundwork for how performance is evaluated and rewards are allocated. For example, in occupations such as offshore drilling or firefighting, overall cultural norms are likely to favor men, and status is likely to be confounded with gender. In these settings, women’s skill attainment may not offset the overall rewards gap because gender remains a dominant status cue and other status signals, such as educational expertise, are often discounted. Clearly, in these situations, women face barriers to translating any human capital gains into greater rewards at work (e.g., Joshi, 2014). We surmise that one important pathway to gender inequality in these settings is the interactional dynamics with peers and supervisors that often drive performance evaluations and reward allocation decisions (see also Joshi & Knight, 2015). We note that an increasing number of organizations are adopting “high involvement” performance management systems that rely on peer evaluations, which are also tied to reward allocation (Cappelli, 1999). These practices place a premium on social interactions among employees, and the extent to which their peers and supervisors recognize and utilize their expertise can be particularly consequential for women’s career advancement opportunities in these contexts (Joshi, 2014; Joshi & Knight, 2015; Ridgeway, 1991). Indeed, scholars have noted that these “strategic,” or “high involvement,” performance management and compensation practices serve to enhance, rather than reduce, gender inequality in organizations (e.g., Cappelli, 1999; Castilla, 2012). We join these scholars in calling for more research on how these high-involvement HR practices serve as

TABLE 4
Results of Qualitative Comparative Analysis

	Panel 1 Men receive higher performance evaluations than women	Panel 2 Women receive higher performance evaluations than men	Panel 3 Men receive higher rewards than women
	1	1 2	1
<i>Occupational context</i>			
Balanced occupation	⊗	●	⊗
High prestige occupation	⊗		●
<i>Industry context</i>			
High proportion of women in executive or senior manager levels	⊗	⊗	●
<i>Job context</i>			
High job complexity		⊗	⊗
Consistency	1.00	1.00	1.00
Raw coverage	0.25	0.20	0.20
Unique coverage	0.25	0.20	0.20
Overall solution consistency	1.00	1.00	1.00
Overall solution coverage	0.25	0.40	0.14

Note: Central conditions are represented by ● (presence) and ⊗ (absence); peripheral conditions are represented by ● (presence) and ⊗ (absence); threshold for consistency was 1.00; frequency was three cases/configuration; no configurations were sufficient for observing women receiving equal or higher rewards than men

mechanisms driving gender inequality and on whether these practices have different outcomes in traditionally male-dominated versus gender-integrated settings.

A particularly striking pattern in our findings is that, in highly prestigious occupations, women did not perform at lower levels than men, and yet men were rewarded significantly higher than women. We propose that, in these settings, it is particularly important to disentangle reward allocation decisions from performance evaluation decisions. In prestigious settings, employment practices may serve as hierarchy-enhancing agents that maintain social hierarchies facilitating male dominance (Sidanius et al., 1996). Indeed, professions such as academia and law offer particularly interesting avenues for further inquiry: In these high-prestige settings, performance criteria tend to be objective (e.g., billable hours or research productivity), yet reward allocation decision making is highly subjective, opaque, and adversarial, and often involves high stakes (many such settings having “up or out” promotion norms). We surmise that this discrepancy in performance evaluations and reward allocation decision making in prestigious contexts is an important mechanism explaining inequity in rewards. Future research could shed light on whether the adoption of promotion practices that do not involve zero-sum choices (e.g., law firms that offer nonequity partnerships) attenuates the social dominance orientation of key decision makers, and related

gender-based attributions of skills and competence, during performance or promotion reviews.

Our findings regarding the effects of job complexity and the sex composition of the upper echelons at the industry level also provide additional avenues for future research. We found that job complexity enhances sex differences in rewards and performance because complex jobs often entail ambiguous performance criteria and idiosyncratic roles that are less open to scrutiny. We propose that future research further unpack how job complexity, via mechanisms such as social comparisons and scrutiny among incumbents, predicts sex differences in employment outcomes. Reinforcing our meta-regression results, our QCA analyses revealed that both configurations representing situations in which women received higher evaluations than men involved less complex jobs (e.g., bank tellers). These routine jobs allow women to select shifts that are conducive to maintaining a work–life balance, are less likely to involve variable time-related demands, and do not place value on putting in “face time” outside of work hours. Even when the occupation was male-dominated or the industry-level executive ranks were predominantly male, our data revealed that women received higher performance evaluations than men while performing routine jobs. These findings underscore the importance of job structure as a mechanism explaining gender inequality, and call for more research into how job design can mitigate sex

differences in performance evaluations, bonus allocations, and promotion rates (see also Correll et al., 2014).

Our findings regarding the effects of the proportion of female executives in specific industry categories further highlight the importance of scrutiny and monitoring for reducing gender inequality. We found that women were rewarded at higher levels than men *and* received higher performance evaluations in only one setting: industries with a high proportion of female executives. The industry environment—specifically, the proportion of female executives—reflects institutional pressures driving the diffusion of egalitarian practices that also support greater monitoring of wage-setting and performance-related practices within firms (Dobbin, 2009). Thus the potential for scrutiny within firms may be driven by these industry-level factors, with important implications for sex differences in promotion and turnover rates within firms. Apart from the potential for scrutiny and monitoring, the proportion of women in executive positions at the industry level also represents greater access to social capital among women. For women, this access to industry-level networks can also help to identify alternative job opportunities within the industry, thereby improving the potential to negotiate a favorable employment contract. While past research has primarily focused on the implications of women's exclusion from social networks within organizations (e.g., Ibarra, 1992), we propose that future studies also examine whether the availability of extraorganizational networks facilitates women's advancement within firms across industry settings.

Caveats and Limitations

Our meta-analysis provides a comprehensive overview of sex differences in employment outcomes during a 30-year period and presents unique insights into the effects of various contextual factors influencing gender equity within firms—but it also has several limitations. First, although the overall sample size was fairly large, the meta-regression analyses involved relatively small samples. Our sample was restricted because some studies in our sample did not provide specific information to code, and we excluded several studies that involved multiple, very diverse work settings from the meta-regression analyses. Only studies that had exact information to match external data sources (e.g., BLS datasets, *O*Net*) were included in our final analyses. Owing to

these restrictions, we could not test an overall model including all of our key and control variables at once.

Second, we were also limited in testing the influence of other potentially important contextual variables on gender-based inequity in work outcomes. For example, organization-level variables such as culture, business strategy, or HR policies have an important influence on sex differences in performance and rewards, because they can shape employees' gender-based attributions and the discretionary behaviors of managers (see Gelfand, Raver, Nishii, & Schneider, 2005, for a review). Moreover, new structuralist perspectives in sociology have examined specific organizational features such as the formalization of procedures (e.g., Elvira & Graham, 2002), implementation of merit-based pay plans (Castilla, 2012), and restructuring (Dencker, 2008) in relation to sex differences in rewards. These studies, unlike ours, focus on single organizational contexts and provide more granular insights into mechanisms governing gender inequality. We could not incorporate these variables in the present model because of the lack of consistent information across the primary studies. Certainly, one avenue for future research may be to examine the effects of these organizational attributes on sex differences in performance or rewards within a broader context, such as in a highly prestigious occupation or in a specific industry context.

Another important limitation of our study is that, given our meta-analytic approach, we were unable to include fine-grained human capital controls such as labor market experience or educational level of the sample. Although we find that sex differences in rewards persisted even after accounting for performance differences and we controlled for the overall tenure of the sample, this is a critical gap. We note that the duration of studies included in our meta-analysis involves a period in which women have more than closed the gap in education and skills attainment. We therefore hope that our findings provide additional explanations for continued gender inequality in the workplace.

On a related note, we were unable to isolate “selection effects” or ascertain what types of process led men and women to choose specific jobs or occupations. For example, research shows that women often opt out of certain jobs even though they perform them as well as men because they differ from many men in their risk-seeking, competitive, and social preferences (Croson & Gneezy, 2009; Niederle & Vesterlund, 2007). It is therefore likely that the women who do choose to stay in high-prestige occupations or complex jobs are markedly different

from others based on some unmeasured attribute(s). We were not able to account for these variables characterizing men and women in our sample. We propose that future meta-analyses compare effect sizes from studies that generate selection-biased estimates (i.e., based on ordinary least squares, or OLS, analysis) with those from selection-corrected models (e.g., models using instrumental variables) across various contexts to examine how these estimates may vary. Additionally, future research may extend the argument developed here to other types of work outcome, such as contextual performance or creativity, or withdrawal behaviors such as turnover, and also examine whether sex differences vary for specific types of reward outcome (e.g., salary increases versus bonuses or promotions). We could not include those variables in the current study because of data limitations, as well as theoretical parsimony.

Finally, while the examination of our funnel plots did not reveal serious publication bias issues, we acknowledge that—especially on a topic such as gender—sometimes not only do null results get filed away, but also politically incorrect results (e.g., “women consistently obtain more rewards than men”). This tendency may yield biased estimates in the sample and cause a potential “file drawer” problem for the meta-analysis.

Implications for Diversity Management

Despite the growing adoption of diversity management practices, research has found that many of these practices (such as mentoring, network groups, or diversity training) have no direct benefits for enhancing access for underrepresented demographic groups within firms and, in fact, that some diversity management practices may even be detrimental to the advancement of underrepresented groups (Kalev et al., 2006). Our findings offer some practical insights that might mitigate gender-based inequity across various work settings and help to tailor diversity management practices. Broadly, we propose that diversity management practices focus on three issues: integrating accountability structures into performance management and compensation practices, designing jobs to promote greater equity among incumbents, and implementing industry-wide mentoring programs for women.

It appears that the potential for scrutinizing performance management and reward allocation processes is one mechanism by which various contextual factors may influence sex differences

in rewards or performance. We propose that accountability structures be built into performance management and reward allocation processes to formalize the scrutiny of these practices, particularly in contexts that are more susceptible to gender inequality. Research shows that the presence of compliance or affirmative action officers has positive effects on the hiring and promotion of women and minorities in organizations (e.g., Edelman, 1992). These findings may be extended to other practices as well. Although it may not be possible to make evaluative or compensation-related decision making fully transparent to employees in many work settings, the presence of neutral third-party observers or advisors can introduce checks and balances to avoid any systematic differences based on gender. For example, in many organizations, managers hold “calibration meetings” to allocate performance ratings to all of their direct reports and then allocate bonuses according to these ratings. Trained neutral observers or advisors present at these meetings may be able to direct discussion away from decision making based on stereotypes or biases toward performance-related information.

A second implication of the effects of job context in our study is that organizations seeking to address inequity introduce interdependence among incumbents in job categories that are more complex or ambiguous in their structure. Although task interdependence among employees within jobs is common, outcome interdependence, such as bonuses based on collective performance, may be critical in these contexts. This type of interdependence could motivate coworkers to seek individuating information regarding underrepresented groups, thereby avoiding reliance on stereotype-based cues of expertise or competence (Fiske, 1998). Complex jobs can also be designed to factor in the time spent in after-hours socializing or face time with clients, so that employees with caring responsibilities are not disadvantaged. Consider, for example, the experience of a major professional accounting firm that found that changing the job requirement to be present at the client site from five days a week to three days a week led to greater engagement and lower turnover among women *and* men in the firm, without any detrimental effects on firm revenues (McCracken, 2000).

Finally, our finding regarding the effects of industry-wide female representation at executive levels is also noteworthy. Given that mentoring programs within organizations have been unsuccessful in bringing about women’s advancement (Ibarra, Carter, & Silva, 2010), we propose that industry-wide mentoring

groups supplement ongoing career development among women. These network groups can be a source of mentoring and advice from senior-level women, industry-wide, as well as directing women toward opportunities for further skill and personal leadership development.

CONCLUSION

Although gender inequality is widespread, a close examination of management research on sex differences in performance evaluations and rewards showed that these effects also vary considerably across work contexts. We proposed that these variations offer fruitful avenues for building an actionable theory of gender inequality in organizations. This meta-analysis provides some important heuristics to identify contexts in which sex differences in performance evaluations and rewards are more or less likely. In an era in which gender bias or discrimination is rarely overt, or even intentional, identifying the sources of chronic gender inequality offers a compelling, yet challenging, agenda for management research. We call for further theoretical developments in the management domain to identify how, in specific work contexts, seemingly neutral organizational practices shape the structure of work, the definition and evaluation of performance, and the allocation of rewards to reinforce gender inequality.

REFERENCES

References marked with an asterisk (*) indicate studies included in the meta-analysis.

- *Abraham, J. D., & Morrison, J. D., Jr. 2003. Validity study relationship between the performance perspectives inventory's implementing scale and job performance of incumbents in a dual security/service role. *Applied H.R.M. Research*, 8: 73–76.
- Acker, J. 2006. Inequality regimes: Gender, class, and race in organizations. *Gender & Society*, 20: 441–464.
- *Allen, T. D. 2001. Family supportive work environments: The role of organizational perceptions. *Journal of Vocational Behavior*, 58: 414–435.
- *Allen, T. D. 2006. Rewarding good citizens: The relationship between citizenship behavior, gender, and organizational rewards. *Journal of Applied Social Psychology*, 36: 120–143.
- *Ang, S., Van Dyne, L., & Begley, T. M. 2003. The employment relationships of foreign workers versus local employees: A field study of organizational justice, job satisfaction, performance, and OCB. *Journal of Organizational Behavior*, 24: 561–583.
- *Aquino, K., & Bommer, W. H. 2003. Preferential mistreatment: How victim status moderates the relationship between organizational citizenship behavior and workplace victimization. *Organization Science*, 14: 374–385.
- *Arvey, R. D., Miller, H. E., Gould, R., & Burch, P. 1987. Interview validity for selecting sales clerks. *Personnel Psychology*, 40: 1–12.
- Auster, E. R. 1989. Task characteristics as a bridge between macro- and micro-level research on salary inequality between men and women. *Academy of Management Review*, 14: 173–193.
- Baron, J. N., & Pfeffer, J. 1994. The social psychology of organizations and inequality. *Social Psychology Quarterly*, 57: 190–209.
- *Barsness, Z. I., Diekmann, K. A., & Seidel, M. D. L. 2005. Motivation and opportunity: The role of remote work, demographic dissimilarity, and social network centrality in impression management. *Academy of Management Journal*, 48: 401–419.
- Bartol, K. 1999. Gender influences on performance evaluations. In G. Powell (Ed.), *Handbook of gender and work*: 165–178. Thousand Oaks, CA: Sage.
- *Bergeron, D. M., Shipp, A. J., Rosen, B., & Furst, S. A. 2013. Organizational citizenship behavior and career outcomes: The cost of being a good citizen. *Journal of Management*, 39: 958–984.
- Bergmann, B. 1974. Occupational segregation, wages, and profits when employers discriminate by race or sex. *Eastern Economic Journal*, 1: 103–110.
- Bernhardt, A., Morris, M., & Handcock, M. 1995. Women's gains or men's losses? A closer look at the shrinking gender gap in earnings. *American Journal of Sociology*, 101: 302–328.
- *Bertrand, M., & Hallock, K. F. 2001. The gender gap in top corporate jobs. *Industrial & Labor Relations Review*, 55: 3–21.
- *Bhatnagar, J. 2014. Mediator analysis in the management of innovation in Indian knowledge workers: The role of perceived supervisor support, psychological contract, reward and recognition and turnover intention. *International Journal of Human Resource Management*, 25: 1395–1416.
- *Bidwell, M. 2011. Paying more to get less: The effects of external hiring versus internal mobility. *Administrative Science Quarterly*, 56: 369–407.
- *Bizzi, L. 2013. The dark side of structural holes: A multilevel investigation. *Journal of Management*, Advance online publication. [DOI: 10.1177/0149206312471385].
- Blau, F. D., & Devaro, J. 2007. New evidence on gender differences in promotion rates: An empirical analysis of a sample of new hires. *Industrial Relations*, 46: 511–550.

- Blau, F. D., & Kahn, L. M. 1981. Race and sex differences in quits by young workers. *Industrial & Labor Relations Review*, 34: 563–577.
- Blau, F. D., & Kahn, L. M. 2007. The gender pay gap: Have women gone as far as they can? *The Academy of Management Perspectives*, 21: 7–23.
- *Bolino, M. C., & Turnley, W. H. 2003. Counternormative impression management, likeability, and performance ratings: The use of intimidation in an organizational setting. *Journal of Organizational Behavior*, 24: 237–250.
- Borman, W. C., & Motowidlo, S. J. 1997. Task performance and contextual performance: The meaning for personnel selection research. *Human Performance*, 10: 99–118.
- Bose, C., & Rossi, P. 1983. Gender and jobs: Prestige standings of occupations as affected by gender. *American Sociological Review*, 48: 316–330.
- *Brass, D. J. 1985. Men's and women's networks: A study of interaction patterns and influence in an organization. *Academy of Management Journal*, 28: 327–343.
- Briscoe, F., & von Nordenflycht, A. 2014. Which path to power? Workplace networks and the relative effectiveness of inheritance and rainmaking strategies for professional partners. *Journal of Professions and Organization*, 1: 33–48.
- *Byrne, Z. S. 2005. Fairness reduces the negative effects of organizational politics on turnover intentions, citizenship behavior and job performance. *Journal of Business and Psychology*, 20: 175–200.
- *Byron, K. 2007. Male and female managers' ability to read emotions: Relationships with supervisor's performance ratings and subordinates' satisfaction ratings. *Journal of Occupational and Organizational Psychology*, 80: 713–733.
- Byron, K., & Khazanchi, S. 2012. Rewards and creative performance: A meta-analytic test of theoretically derived hypotheses. *Psychological Bulletin*, 138: 809–830.
- Campbell, D. J. 1988. Task complexity: A review and analysis. *Academy of Management Review*, 13: 40–52.
- Cappelli, P. 1999. *The new deal at work*. Boston, MA: Harvard Business School Press.
- *Carless, S. A., & Arnup, J. L. 2011. A longitudinal study of the determinants and outcomes of career change. *Journal of Vocational Behavior*, 78: 80–91.
- *Carlson, D. S., Ferguson, M., Kacmar, K. M., Grzywacz, J. G., & Whitten, D. 2011. Pay it forward: The positive crossover effects of supervisor work–family enrichment. *Journal of Management*, 37: 770–789.
- *Carmeli, A., & Josman, Z. E. 2006. The relationship among emotional intelligence, task performance, and organizational citizenship behaviors. *Human Performance*, 19: 403–419.
- Carney, M., Gedajlovic, E. R., Heugens, P. P., Van Essen, M., & Van Oosterhout, J. H. 2011. Business group affiliation, performance, context, and strategy: A meta-analysis. *Academy of Management Journal*, 54: 437–460.
- Castilla, E. J. 2008. Gender, race, and meritocracy in organizational careers. *American Journal of Sociology*, 113: 1479–1526.
- *Castilla, E. J. 2012. Gender, race, and the new (merit-based) employment relationship. *Industrial Relations*, 51: 528–562.
- Castilla, E. J., & Benard, S. 2010. The paradox of meritocracy in organizations. *Administrative Science Quarterly*, 55: 543–576.
- Catalyst. 2008. Catalyst 2008 census of the Fortune 500 reveals women gained little ground advancing to business leadership positions. Accessed online at <http://www.catalyst.org/press-release/141/catalyst-2008-census-of-the-fortune-500-reveals-women-gained-little-ground-advancing-to-business-leadership-positions>.
- *Chatman, J. A., Polzer, J. T., Barsade, S. G., & Neale, M. A. 1998. Being different yet feeling similar: The influence of demographic composition and organizational culture on work processes and outcomes. *Administrative Science Quarterly*, 43: 749–780.
- *Chauvin, K. W., & Ash, R. A. 1994. Gender earnings differentials in total pay, base pay, and contingent pay. *Industrial & Labor Relations Review*, 47: 634–649.
- *Chen, Z. X., & Aryee, S. 2007. Delegation and employee work outcomes: An examination of the cultural context of mediating processes in China. *Academy of Management Journal*, 50: 226–238.
- *Chen, Z. X., & Francesco, A. M. 2003. The relationship between the three components of commitment and employee performance in China. *Journal of Vocational Behavior*, 62: 490–510.
- *Chen, Z. X., Lam, W., & Zhong, J. A. 2007. Leader-member exchange and member performance: A new look at individual-level negative feedback-seeking behavior and team-level empowerment climate. *The Journal of Applied Psychology*, 92: 202–219.
- *Chen, Z. X., Tsui, A. S., & Farh, J. L. 2002. Loyalty to supervisor vs. organizational commitment: Relationships to employee performance in China. *Journal of Occupational and Organizational Psychology*, 75: 339–356.
- *Chou, L. F., Cheng, B. S., Huang, M. P., & Cheng, H. Y. 2006. Guanxi networks and members' effectiveness in Chinese work teams: Mediating effects of trust networks. *Asian Journal of Social Psychology*, 9: 79–95.

- Cleveland, J. N., Vescio, T. K., & Barnes-Farrell, J. L. 2005. Gender discrimination in organizations. In R. L. Dipboye & A. Colella (Eds.), *Discrimination at work: The psychological and organizational bases*: 149–176. Mahwah, NJ: Lawrence Erlbaum Associates.
- Cohen, L. E., & Broschak, J. P. 2013. Whose jobs are these? The impact of the proportion of female managers on the number of new management jobs filled by women versus men. *Administrative Science Quarterly*, 2: 1–33.
- Cohen, L. E., Broschak, J. P., & Haveman, H. A. 1998. And then there were more? The effect of organizational sex composition on the hiring and promotion of managers. *American Sociological Review*, 63: 711–727.
- *Colquitt, J. A., LePine, J. A., Zapata, C. P., & Wild, R. E. 2011. Trust in typical and high-reliability contexts: Building and reacting to trust among firefighters. *Academy of Management Journal*, 54: 999–1015.
- Correll, S., Kelly, E., O'Connor, L. T., & Williams, J. 2014. Redesigning, redefining work. *Work and Occupations*, 41: 3–17.
- *Cropanzano, R., Rupp, D. E., & Byrne, Z. S. 2003. The relationship of emotional exhaustion to work attitudes, job performance, and organizational citizenship behaviors. *The Journal of Applied Psychology*, 88: 160–169.
- Croson, R., & Gneezy, U. 2009. Gender differences in preferences. *Journal of Economic Literature*, 47: 448–474.
- *Day, R., & Allen, T. D. 2004. The relationship between career motivation and self-efficacy with protégé career success. *Journal of Vocational Behavior*, 64: 72–91.
- Dencker, J. C. 2008. Corporate restructuring and sex differences in managerial promotion. *American Sociological Review*, 73: 455–476.
- *Diefendorff, J. M., Brown, D. J., Kamin, A. M., & Lord, R. G. 2002. Examining the roles of job involvement and work centrality in predicting organizational citizenship behaviors and job performance. *Journal of Organizational Behavior*, 23: 93–108.
- Dipboye, R. L., & Colella, A. 2005. *Discrimination at work: The psychological and organizational bases*. Mahwah, NJ: Lawrence Erlbaum Associates.
- DiTomaso, N., Post, C., & Parks-Yancy, R. 2007a. Workforce diversity and inequality: Power, status, and numbers. *Annual Review of Sociology*, 33: 473–501.
- DiTomaso, N., Post, C., Smith, D. R., Farris, G. F., & Cordero, R. 2007b. Effects of structural position on allocation and evaluation decisions for scientists and engineers in industrial R&D. *Administrative Science Quarterly*, 52: 175–207.
- Dobbin, F. 2009. *Inventing equal opportunity*. Princeton, NJ: Princeton University Press.
- *Dreher, G. F., & Ash, R. A. 1990. A comparative study of mentoring among men and women in managerial, professional, and technical positions. *The Journal of Applied Psychology*, 75: 539–546.
- *Dreher, G. F., & Cox, T. H. 2000. Labor market mobility and cash compensation: The moderating effects of race and gender. *Academy of Management Journal*, 43: 890–900.
- Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. 1992. Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin*, 111: 3–22.
- Eagly, A. H., & Karau, S. J. 1991. Gender and the emergence of leaders: A metaanalysis. *Journal of Personality and Social Psychology*, 60: 685–710.
- *Eby, L. T., Allen, T. D., & Douthitt, S. S. 1999. The role of nonperformance factors on job-related relocation opportunities: A field study and laboratory experiment. *Organizational Behavior and Human Decision Processes*, 79: 29–55.
- Edelman, L. B. 1992. Legal ambiguity and symbolic structures: Organizational mediation of civil rights law. *American Journal of Sociology*, 97: 1531–1576.
- Edelman, L. B., Fuller, S.F., & Mara-Drita, I. 2001. Diversity rhetoric and the managerialization of the law. *American Journal of Sociology*, 106: 1589–1641.
- EEOC. 2002. Job patterns for women and minorities in private industry. <http://www.eeoc.gov/stats/jobpat/jobpat.html>
- *Elfenbein, H. A., & O'Reilly, C. A. 2007. Fitting in: The effects of relational demography and person–culture fit on group process and performance. *Group & Organization Management*, 32: 109–142.
- *Elvira, M. M., & Cohen, L. E. 2001. Location matters: A cross-level analysis of the effects of organizational sex composition on turnover. *Academy of Management Journal*, 44: 591–605.
- *Elvira, M. M., & Graham, M. E. 2002. Not just a formality: Pay system formalization and sex-related earnings effects. *Organization Science*, 13: 601–617.
- Elvira, M., & Town, R. 2001. The effects of race and worker productivity on performance evaluations. *Industrial Relations: A Journal of Economy and Society*, 40: 571–590.
- Ely, R. J. 1994. The effects of organizational demographics and social identity on relationships among professional women. *Administrative Science Quarterly*, 39: 203–238.
- Ely, R. J. 1995. The power in demography: Women's social constructions of gender identity at work. *Academy of Management Journal*, 38: 589–634.

- Ely, R., & Padavic, I. 2007. A feminist analysis of organizational research on sex differences. *Academy of Management Review*, 32: 1121–1143.
- England, P. 1979. Women and occupational prestige: A case of vacuous sex equality. *Signs*, 5: 252–265.
- England, P., & McLaughlin, S. 1979. Sex segregation of jobs and male–female income differentials. In R. Alvarez & K. Lutterman (Eds.), *Associates: Discrimination in organizations*: 189–213. San Francisco, CA: Jossey-Bass.
- *Farh, C. I. C. C., Seo, M. G., & Tesluk, P. E. 2012. Emotional intelligence, teamwork effectiveness, and job performance: The moderating role of job context. *The Journal of Applied Psychology*, 97: 890–900.
- *Feldman, D. C., & Turnley, W. H. 2004. Contingent employment in academic careers: relative deprivation among adjunct faculty. *Journal of Vocational Behavior*, 64: 284–307.
- Felstiner, W., Abel, R., & Sarat, A. 1980. The emergence and transformation of disputes: Naming, blaming, claiming. *Law & Society Review*, 15: 631–654.
- *Ferris, G. R., Bergin, T. G., & Wayne, S. J. 1988. Personal characteristics, job performance, and absenteeism of public school teachers. *Journal of Applied Social Psychology*, 18: 552–563.
- *Ferris, G. R., Witt, L. A., & Hochwarter, W. A. 2001. Interaction of social skill and general mental ability on job performance and salary. *The Journal of Applied Psychology*, 86: 1075–1082.
- *Fields, D., Pang, M., & Chiu, C. 2000. Distributive and procedural justice as predictors of employee outcomes in Hong Kong. *Journal of Organizational Behavior*, 21: 547–562.
- Fiske, A. P. 1998. Human sociality. *International Society for the Study of Personal Relationships Bulletin*, 14: 4–9.
- Fiske, S. T., Bersoff, D. N., Borgida, E., Deaux, K., & Heilman, M. E. 1991. Social science research on trial: Use of sex stereotyping research in *Price Waterhouse v. Hopkins*. *The American Psychologist*, 46: 1049–1060.
- Fiss, P. C. 2011. Building better causal theories: A fuzzy set approach to typologies in organizational research. *Academy of Management Journal*, 54: 393–420.
- *Flynn, F. J., & Ames, D. R. 2006. What's good for the goose may not be as good for the gander: The benefits of self-monitoring for men and women in task groups and dyadic conflicts. *The Journal of Applied Psychology*, 91: 272–281.
- Gajendran, R. S., & Harrison, D. A. 2007. The good, the bad, and the unknown about telecommuting: Meta-analysis of psychological mediators and individual consequences. *The Journal of Applied Psychology*, 92: 1524.
- Gatta, M. L., & Roos, P. A. 2005. Rethinking occupational integration. *Sociological Forum*, 20: 369–402.
- Gelfand, M., Raver, J., Nishii, L., & Schneider, B. 2005. Discrimination in organizations: An organizational-level systems perspective. In R. Dipboye & A. Colella (Eds.), *Discrimination at work: The psychological and organizational bases*: 89–116. Mahwah, NJ: Lawrence Erlbaum Associates.
- Gerdes, E. P., & Garber, D. M. 1983. Sex bias in hiring: Effects of job demands and applicant competence. *Sex Roles*, 9: 307–319.
- *Gomez-Mejia, L. R., & Balkin, D. B. 1984. Faculty satisfaction with pay and other job dimensions under union and nonunion conditions. *Academy of Management Journal*, 27: 591–602.
- *Gomez-Mejia, L. R., & Balkin, D. B. 1992. Determinants of faculty pay: An agency theory perspective. *Academy of Management Journal*, 35: 921–955.
- Gorman, E. H. 2005. Gender stereotypes, same-gender preferences, and organizational variation in the hiring of women: Evidence from law firms. *American Sociological Review*, 70: 702–728.
- Gorman, E. H. 2006. Work uncertainty and the promotion of professional women: The case of law firm partnership. *Social Forces*, 85: 865–890.
- *Gorman, E. H., & Kmec, J. A. 2009. Hierarchical rank and women's organizational mobility: Glass ceilings in corporate law firms. *American Journal of Sociology*, 114: 1428–1474.
- *Graves, L. M., Ruderman, M. N., Ohlott, P. J., & Weber, T. J. 2012. Driven to work and enjoyment of work effects on managers' outcomes. *Journal of Management*, 38: 1655–1680.
- Greckhamer, T., Misangyi, V. F., Elms, H., & Lacey, R. 2008. Using qualitative comparative analysis in management research: An examination of combinations of industry, corporate and business-unit effects. *Organizational Research Methods*, 11: 695–726.
- Greenhaus, J. H., & Parasuraman, S. 1993. Job performance attributions and career advancement prospects: An examination of gender and race effects. *Organizational Behavior and Human Decision Processes*, 55: 273–297.
- *Greenhaus, J. H., Parasuraman, S., & Wormley, W. M. 1990. Effects of race on organizational experiences, job performance evaluations, and career outcomes. *Academy of Management Journal*, 33: 64–86.
- Harbord, R., & Higgins, J. 2008. Meta-regression in STATA. *The Stata Journal*, 8: 492–519.

- Hauser, R. M., & Warren, J. R. 1997. Socioeconomic indexes for occupations: A review, update, and critique. *Sociological Methodology*, 27: 177–298.
- Hedges, L. V., & Olkin, I. 1985. *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- Heilman, M. E. 1983. Sex bias in work settings: The lack of fit model. *Research in Organizational Behavior*, 5: 269–298.
- Heilman, M. E. 1994. Affirmative action: Some unintended consequences. *Research in Organizational Behavior*, 16: 125–170.
- *Hekman, D. R., Aquino, K., Owens, B. P., Mitchell, T. R., Schilpzand, P., & Leavitt, K. 2010. An examination of whether and how racial and gender biases influence customer satisfaction. *Academy of Management Journal*, 53: 238–264.
- *Hersch, J., & Viscusi, W. 1996. Gender differences in promotions and wages. *Industrial Relations: A Journal of Economy and Society*, 35: 461–472.
- Higgins, J., & Thompson, S. 2004. Controlling the risk of spurious findings from meta-regression. *Statistics in Medicine*, 23: 1663–1682.
- Hirsch, C. E., & Kornrich, S. 2008. The context of discrimination: Workplace conditions, institutional environments, and sex and race discrimination charges. *American Journal of Sociology*, 113: 1394–1432.
- Hodge, R. W., Siegel, P. M., & Rossi, P. H. 1964. Occupational prestige in the United States, 1925–63. *American Journal of Sociology*, 70: 286–302.
- *Honeycutt, T. L., & Rosen, B. 1997. Family friendly human resource policies, salary levels, and salient identity as predictors of organizational attraction. *Journal of Vocational Behavior*, 50: 271–290.
- *Hoobler, J. M., Lemmon, G., & Wayne, S. J. 2014. Women's managerial aspirations: An organizational development perspective. *Journal of Management*, 40: 703–730.
- *Hoobler, J. M., Wayne, S. J., & Lemmon, G. 2009. Bosses' perceptions of family-work conflict and women's promotability: Glass ceiling effects. *Academy of Management Journal*, 52: 939–957.
- Huffman, M. L., Cohen, P. N., & Pearlman, J. 2010. Engendering change: Organizational dynamics and workplace gender desegregation, 1975–2005. *Administrative Science Quarterly*, 55: 255–277.
- *Hui, C., Lam, S. S., & Law, K. K. 2000. Instrumental values of organizational citizenship behavior for promotion: A field quasi-experiment. *The Journal of Applied Psychology*, 85: 822–828.
- Hultin, M., & Szulkin, R. 2003. Mechanisms of inequality: Unequal access to organizational power and the gender wage gap. *European Sociological Review*, 19: 143–159.
- Ibarra, H. 1992. Homophily and differential returns: Sex differences in network structure and access in an advertising firm. *Administrative Science Quarterly*, 37: 422–447.
- Ibarra, H., Carter, N., & Silva, C. 2010. Why men still get more promotions than women. *Harvard Business Review*, 88: 80–85.
- *Igbaria, M., & Baroudi, J. 1995. The impact of job performance evaluation on career advancement prospects: An examination of gender differences in the workplace. *Management Information Systems Quarterly*, 19: 5–15.
- *Indartono, S., & Chen, C. V. 2010. Moderation of gender on the relationship between task characteristics and performance. *International Journal of Organizational Innovation*, 2: 195–200.
- *Iverson, R. D., & Kuruvilla, S. 1995. Antecedents of union loyalty: The influence of individual dispositions and organizational context. *Journal of Organizational Behavior*, 16: 557–582.
- *Iverson, R. D., & Pullman, J. A. 2000. Determinants of voluntary turnover and layoffs in an environment of repeated downsizing following a merger: An event history analysis. *Journal of Management*, 26: 977–1003.
- *James, E. H. 2000. Race-related differences in promotions and support: Underlying effects of human and social capital. *Organization Science*, 11: 493–508.
- *Jepsen, D. M., & Rodwell, J. J. 2009. Justice in the workplace: The centrality of social versus judgmental predictors of performance varies by gender. *International Journal of Human Resource Management*, 20: 2066–2083.
- Joshi, A. 2014. By whom and when is women's expertise recognized? The interactive effects of gender and education in science and engineering teams. *Administrative Science Quarterly*, 59: 202–239.
- Joshi, A., & Knight, A. 2015. Who defers to whom and why? Dual mechanisms linking demographic differences to dyadic deference and team effectiveness. *Academy of Management Journal*, 58: 59–84.
- *Joshi, A., Liao, H., & Jackson, S. E. 2006. Cross-level effects of workplace diversity on sales performance and pay. *Academy of Management Journal*, 49: 459–481.
- *Judge, T. A., & Cable, D. M. 2011. When it comes to pay, do the thin win? The effect of weight on pay for men and women. *The Journal of Applied Psychology*, 96: 95.
- *Judge, T. A., & Watanabe, S. 1993. Another look at the job satisfaction–life satisfaction relationship. *The Journal of Applied Psychology*, 78: 939–948.

- *Judge, T. A., Cable, D. M., Boudreau, J. W., & Bretz, R. D. 1995. An empirical investigation of the predictors of executive career success. *Personnel Psychology*, 48: 485–519.
- *Judiesch, M. K., & Lyness, K. S. 1999. Left behind? The impact of leaves of absence on managers' career success. *Academy of Management Journal*, 42: 641–651.
- Kalev, A., Dobbin, F., & Kelly, E. 2006. Best practices or best guesses? Assessing the efficacy of corporate affirmative action and diversity policies. *American Sociological Review*, 71: 589–617.
- *Kamdar, D., & Van Dyne, L. 2007. The joint effects of personality and workplace social exchange relationships in predicting task performance and citizenship performance. *The Journal of Applied Psychology*, 92: 1286.
- Kanter, R. M. 1977. *Men and women of the corporation*. New York: Basic Books.
- Kay, F. M., & Hagan, J. 1998. Raising the bar: The gender stratification of law-firm capital. *American Sociological Review*, 64: 728–743.
- *Keller, R. T. 2012. Predicting the performance and innovativeness of scientists and engineers. *The Journal of Applied Psychology*, 97: 225–233.
- Kepes, S., Banks, G. C., McDaniel, M., & Whetzel, D. L. 2012. Publication bias in the organizational sciences. *Organizational Research Methods*, 15: 624–662.
- *Kirchmeyer, C. 1995. Demographic similarity to the work group: A longitudinal study of managers at the early career stage. *Journal of Organizational Behavior*, 16: 67–83.
- *Kirchmeyer, C. 1998. Determinants of managerial career success: Evidence and explanation of male/female differences. *Journal of Management*, 24: 673–692.
- *Kirchmeyer, C. 2002. Gender differences in managerial careers: Yesterday, today, and tomorrow. *Journal of Business Ethics*, 37: 5–24.
- *Kluemper, D. H., DeGroot, T., & Choi, S. 2013. Emotion management ability predicting task performance, citizenship, and deviance. *Journal of Management*, 39: 878–905.
- *Lam, S. S., Ng, T. W., & Feldman, D. C. 2012. The relationship between external job mobility and salary attainment across career stages. *Journal of Vocational Behavior*, 80: 129–136.
- *Landau, J. 1995. The relationship of race and gender to managers' ratings of promotion potential. *Journal of Organizational Behavior*, 16: 391–400.
- Lansberg, I. 1989. Social categorization, entitlement, and justice in organizations: Contextual determinants and cognitive underpinnings. *Human Relations*, 41: 871–899.
- Lewis-Beck, M. S., Bryman, A. E., & Liao, T. F. 2004. *The Sage encyclopedia of social science research methods, vol. 1*. Thousand Oaks, CA: Sage.
- Lipsey, M. W., & Wilson, D. B. 2001. *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- *Liu, J., Hui, C., Lee, C., & Chen, Z. X. 2013. Why do I feel valued and why do I contribute? A relational approach to employee's organization-based self-esteem and job performance. *Journal of Management Studies*, 50: 1018–1040.
- *Liu, Y., Loi, R., & Lam, L. W. 2011. Linking organizational identification and employee performance in teams: The moderating role of team-member exchange. *International Journal of Human Resource Management*, 22: 3187–3201.
- *Lobel, S. A., & St. Clair, L. 1992. Effects of family responsibilities, gender, and career identity salience on performance outcomes. *Academy of Management Journal*, 35: 1057–1069.
- *Loi, R., & Ngo, H. Y. 2009. Work outcomes of relational demography in Chinese vertical dyads. *International Journal of Human Resource Management*, 20: 1704–1719.
- *Lyness, K. S., & Heilman, M. E. 2006. When fit is fundamental: Performance evaluations and promotions of upper-level female and male managers. *The Journal of Applied Psychology*, 91: 777–785.
- *Lyness, K. S., & Judiesch, M. K. 1999. Are women more likely to be hired or promoted into management positions? *Journal of Vocational Behavior*, 54: 158–173.
- *Lyness, K. S., & Judiesch, M. K. 2001. Are female managers quitters? The relationships of gender, promotions, and family leaves of absence to voluntary turnover. *The Journal of Applied Psychology*, 86: 1167–1178.
- *Lyness, K. S., & Thompson, D. E. 1997. Above the glass ceiling? A comparison of matched samples of female and male executives. *The Journal of Applied Psychology*, 82: 359–375.
- Martin, J., & Myerson, D. D. 1998. Women and power: Conformity, resistance, and disorganized coercion. In R. Kramer & M. Neale (Eds.), *Power and influence in organizations*: 311–348. Thousand Oaks, CA: Sage.
- *Martins, L. L., Eddleston, K. A., & Veiga, J. F. 2002. Moderators of the relationship between work-family conflict and career satisfaction. *Academy of Management Journal*, 45: 399–409.
- McCall, L. 2001. *Complex inequality: Gender, class, and race in the new economy*. New York: Routledge.

- McCracken, D. M. 2000. Winning the talent war on women: Sometimes it takes a revolution. *Harvard Business Review*, 78: 159–167.
- McGahan, A. M., & Porter, M. E. 1997. How much does industry matter, really? *Strategic Management Journal*, 18: 15–30.
- *Miller, J. G., & Wheeler, K. G. 1992. Unraveling the mysteries of gender differences in intentions to leave the organization. *Journal of Organizational Behavior*, 13: 465–478.
- Misangyi, V. F., & Acharya, A. 2014. Substitutes or complements? A configurational examination of corporate governance mechanisms. *Academy of Management Journal*, 57: 1681–1705.
- Mobley, W. H. 1982. Supervisor and employee race and sex effects on performance appraisals: A field study of adverse impact and generalizability. *Academy of Management Journal*, 25: 598–606.
- Morrison, A. M., & Von Glinow, M. A. 1990. Women and minorities in management. *The American Psychologist*, 45: 200–208.
- *Morrison, E. W. 1994. Role definitions and organizational citizenship behavior: The importance of the employee's perspective. *Academy of Management Journal*, 37: 1543–1567.
- *Mount, M. K., In-Sue, O., & Burns, M. 2008. Incremental validity of perceptual speed and accuracy over general mental ability. *Personnel Psychology*, 61: 113–139.
- Nakao, K., & Treas, J. 1994. Updating occupational prestige and socioeconomic scores: How the new measures measure up. In P. Marsden (Ed.), *Sociological methodology*: 1–72, Washington, DC: American Sociological Association.
- Nelson, R. R., & Bridges, W. P. 1999. *Legalizing gender inequality: Court, markets, and unequal pay for women in America*. Cambridge: Cambridge University Press.
- Ng, T. W. H., Eby, L. T., Sorensen, K. L., & Feldman, D. C. 2005. Predictors of objective and subjective career success: A meta-analysis. *Personnel Psychology*, 58: 367–408.
- Niederle, M., & Vesterlund, L. 2007. Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics*, 122: 1067–1101.
- Nieva, V. F., & Gutek, B. A. 1980. Sex effects on evaluation. *Academy of Management Review*, 5: 267–276.
- *Nyberg, A. J., Pieper, J. R., & Trevor, C. O. 2013. Pay-for-performance's effect on future employee performance integrating psychological and economic principles toward a contingency perspective. *Journal of Management*, Advance online publication. [DOI: 10.1177/0149206313515520].
- Oritz, S., & Roscigno, V. 2009. Discrimination, women, and work: Processes and variations by race and class. *The Sociological Quarterly*, 50: 336–359.
- Ostroff, C., & Atwater, L. E. 2003. Does whom you work with matter? Effects of referent group gender and age composition on managers' compensation. *The Journal of Applied Psychology*, 88: 725–740.
- *Ostroff, C., Atwater, L. E., & Feinberg, B. J. 2004. Understanding self-other agreement: A look at rater and ratee characteristics, context, and outcomes. *Personnel Psychology*, 57: 333–375.
- *Peterson, S. J., Luthans, F., Avolio, B. J., Walumbwa, F. O., & Zhang, Z. 2011. Psychological capital and employee performance: A latent growth modeling approach. *Personnel Psychology*, 64: 427–450.
- Podsakoff, N. P., Whiting, S. W., Podsakoff, P. M., & Blume, B. D. 2009. Individual- and organizational-level consequences of organizational citizenship behaviors: A meta-analysis. *The Journal of Applied Psychology*, 94: 122–141.
- Porter, M. E. 1980. *Competitive strategy*. New York: Free Press.
- *Powell, G. N., & Butterfield, D. A. 1994. Investigating the “glass ceiling” phenomenon: An empirical study of actual promotions to top management. *Academy of Management Journal*, 37: 68–86.
- Pratto, F., Stallworth, L. M., Sidanius, J., & Siers, B. 1997. The gender gap in occupational role attainment: A social dominance approach. *Journal of Personality and Social Psychology*, 72: 37–53.
- Ragin, C. C. 2000. *Fuzzy-set social science*. Chicago, IL: University of Chicago Press.
- Ragin, C. C. 2008. *Redesigning social inquiry: Fuzzy sets and beyond*. Chicago, IL: University of Chicago Press.
- Ragin, C. C., & Rihoux, B. 2009. Qualitative comparative analysis (QCA) and related systematic comparative methods. *International Sociology*, 21: 679–706.
- *Ragins, B. R., & Cornwell, J. M. 2001. Pink triangles: Antecedents and consequences of perceived workplace discrimination against gay and lesbian employees. *The Journal of Applied Psychology*, 86: 1244–1267.
- *Randel, A. E., & Jaussi, K. S. 2003. Functional background identity, diversity, and individual performance in cross-functional teams. *Academy of Management Journal*, 46: 763–774.
- Reskin, B. F. 2003. Including mechanisms of ascriptive inequality. *American Sociological Review*, 68: 1–21.
- *Restubog, S. L. D., Bordia, P., & Tang, R. L. 2007. Behavioural outcomes of psychological contract breach in a non-Western culture: The moderating role of equity sensitivity. *British Journal of Management*, 18: 376–386.

- *Restubog, S. L. D., Bordia, P., Tang, R. L., & Krebs, S. A. 2010. Investigating the moderating effects of leader-member exchange in the psychological contract breach-employee performance relationship: A test of two competing perspectives. *British Journal of Management*, 21: 422-437.
- Ridgeway, C. 1991. The social construction of status value: Gender and other nominal characteristics. *Social Forces*, 70: 367-386.
- *Rosin, H., & Korabik, K. 1995. Organizational experiences and propensity to leave: A multivariate investigation of men and women managers. *Journal of Vocational Behavior*, 46: 1-16.
- Rudman, L. A., Moss-Racusin, C. A., Phelan, J. E., & Nauts, S. 2012. Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against female leaders. *Journal of Experimental Social Psychology*, 48: 165-179.
- Sackett, P. R., DuBois, C. L., & Noe, A. W. 1991. Tokenism in performance evaluation: The effects of work group representation on male-female and white-black differences in performance ratings. *The Journal of Applied Psychology*, 76: 263-267.
- *Salamin, A., & Hom, P. W. 2005. In search of the elusive U-shaped performance-turnover relationship: Are high-performing Swiss bankers more liable to quit? *The Journal of Applied Psychology*, 90: 1204-1230.
- *Schaubroeck, J., & Lam, S. S. 2002. How similarity to peers and supervisor influences organizational advancement in different cultures. *Academy of Management Journal*, 45: 1120-1136.
- *Schneer, J. A., & Reitman, F. 1990. Effects of employment gaps on the careers of MBAs: More damaging for men than for women? *Academy of Management Journal*, 33: 391-406.
- *Schneer, J. A., & Reitman, F. 1993. Effects of alternate family structures on managerial career paths. *Academy of Management Journal*, 36: 830-843.
- *Schneer, J. A., & Reitman, F. 1994. The importance of gender in mid-career: A longitudinal study of MBAs. *Journal of Organizational Behavior*, 15: 199-207.
- *Seibert, S. E., Grant, J. M., & Kraimer, M. L. 1999. Proactive personality and career success. *The Journal of Applied Psychology*, 84: 416-427.
- *Seibert, S. E., Kraimer, M. L., & Liden, R. C. 2001. A social capital theory of career success. *Academy of Management Journal*, 44: 219-237.
- *Shaw, J. D., Duffy, M. K., Mitra, A., Lockhart, D. E., & Bowler, M. 2003. Reactions to merit pay increases: A longitudinal test of a signal sensitivity perspective. *The Journal of Applied Psychology*, 88: 538-553.
- *Shi, J., Chen, Z., & Zhou, L. 2011. Testing differential mediation effects of sub-dimensions of political skills in linking proactive personality to employee performance. *Journal of Business and Psychology*, 26: 359-369.
- *Shoss, M. K., Eisenberger, R., Restubog, S. L. D., & Zagenczyk, T. J. 2013. Blaming the organization for abusive supervision: The roles of perceived organizational support and supervisor's organizational embodiment. *The Journal of Applied Psychology*, 98: 158-168.
- Sidanius, J., & Pratto, F. 1999. *Social dominance: An intergroup theory of social hierarchy and oppression*. New York: Cambridge University Press.
- Sidanius, J., Pratto, F., Sinclair, S., & van Laar, C. 1996. Mother Teresa meets Genghis Khan: The dialectics of hierarchy-enhancing and hierarchy-attenuating career choices. *Social Justice Research*, 9: 145-170.
- Siegel, P. M. 1971. *Prestige in the American occupational structure*. Unpublished PhD thesis, University of Chicago, Chicago, IL.
- Sinclair, S., Sidanius, J., & Levin, S. 2010. The interface between ethnic and social system attachment: The differential effects of hierarchy-enhancing and hierarchy-attenuating environments. *The Journal of Social Issues*, 54: 741-757.
- *Sosik, J. J., Juzbasich, J., & Chun, J. U. 2011. Effects of moral reasoning and management level on ratings of charismatic leadership, in-role and extra-role performance of managers: A multi-source examination. *The Leadership Quarterly*, 22: 434-450.
- Speilerman, S., & Petersen, T. 1999. Organizational structure, determinants of promotion, and gender differences in attainment. *Social Science Research*, 28: 203-227.
- Spurr, S. J., & Sueyoshi, G. T. 1994. Turnover and promotion of lawyers: An inquiry into gender differences. *The Journal of Human Resources*, 29: 813-842.
- Stainback, K., Tomaskovic-Devey, D., & Skaggs, S. 2010. Organizational approaches to inequality: Inertia, relative power, and environments. *Annual Review of Sociology*, 36: 225-247.
- Stroh, L. K., Brett, J. M., & Reilly, A. H. 1992. All the right stuff: A comparison of female and male managers' career progression. *The Journal of Applied Psychology*, 77: 251-260.
- *Sturges, J., Guest, D., Conway, N., & Mackenzie Davey, K. 2002. A longitudinal study of the relationship between career management and organizational commitment among graduates in the first ten years at work. *Journal of Organizational Behavior*, 23: 731-748.
- *Sturman, M. C., Shao, L., & Katz, J. H. 2012. The effect of culture on the curvilinear relationship between performance and turnover. *The Journal of Applied Psychology*, 97: 46.

- *Tharenou, P. 2001. Going up? Do traits and informal social processes predict advancing in management? *Academy of Management Journal*, 44: 1005–1017.
- Thompson, S. G., & Higgins, J. P. T. 2002. How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine*, 21: 1559–1573.
- Tilly, C. 1998. *Durable inequality*. Berkeley, CA: University of California Press.
- Tomaskovic-Devey, D. 1993. *Gender and racial inequality at work: The sources and consequences of job segregation*. Ithaca, NY: ILR Press.
- Tomaskovic-Devey, D., & Skaggs, S. 1999. An establishment-level test of the statistical discrimination hypothesis. *Work and Occupations*, 26: 422–445.
- *Tremblay, M., St-Onge, S., & Toulouse, J. M. 1997. Determinants of salary referents relevance: A field study of managers. *Journal of Business and Psychology*, 11: 463–484.
- *Trevor, C. O. 2001. Interactions among actual ease-of-movement determinants and job satisfaction in the prediction of voluntary turnover. *Academy of Management Journal*, 44: 621–638.
- *Tsui, A. S., & O'Reilly, C. A., III. 1989. Beyond simple demographic effects: The importance of relational demography in superior-subordinate dyads. *Academy of Management Journal*, 32: 402–423.
- *Turban, D. B., & Dougherty, T. W. 1994. Role of protégé personality in receipt of mentoring and career success. *Academy of Management Journal*, 37: 688–702.
- *Turnley, W. H., & Feldman, D. C. 2000. Re-examining the effects of psychological contract violations: Unmet expectations and job dissatisfaction as mediators. *Journal of Organizational Behavior*, 21: 25–42.
- *Turnley, W. H., Bolino, M. C., Lester, S. W., & Bloodgood, J. M. 2003. The impact of psychological contract fulfillment on the performance of in-role and organizational citizenship behaviors. *Journal of Management*, 29: 187–206.
- *Van Dyne, L., & LePine, J. A. 1998. Helping and voice extra-role behaviors: Evidence of construct and predictive validity. *Academy of Management Journal*, 41: 108–119.
- *Van Iddekinge, C. H., Morgeson, F. P., Schleicher, D. J., & Campion, M. A. 2011. Can I retake it? Exploring subgroup differences and criterion-related validity in promotion retesting. *The Journal of Applied Psychology*, 96: 941.
- Vescio, T. K., Gervais, S., Snyder, M., & Hoover, A. 2005. Power and the creation of patronizing environments: The stereotype-based behaviors of the powerful and their effects on female performance in masculine domains. *Journal of Personality and Social Psychology*, 88: 658–672.
- Viswesvaran, C., & Ones, D. S. 1995. Theory testing: Combining psychometric meta-analysis and structural equations modeling. *Personnel Psychology*, 48: 865–886.
- *Wang, M. O., Zhan, Y., Mccune, E., & Truxillo, D. 2011. Understanding newcomers' adaptability and work-related outcomes: Testing the mediating roles of perceived P–E fit variables. *Personnel Psychology*, 64: 163–189.
- *Watkins, M. B., et al. 2006. Does it pay to be a sexist? The relationship between modern sexism and career outcomes. *Journal of Vocational Behavior*, 69: 524–537.
- Wayne, S. J., & Liden, R. C. 1995. Effects of impression management on performance ratings: A longitudinal study. *Academy of Management Journal*, 38: 232–260.
- Weeden, K. A. 2002. Why some occupations pay more than others? Social closure and earnings inequality in United States. *American Journal of Sociology*, 108: 55–101.
- *Weng, Q., & McElroy, J. C. 2012. Organizational career growth, affective occupational commitment and turnover intentions. *Journal of Vocational Behavior*, 80: 256–265.
- *Whitely, W. T., & Coetsier, P. 1993. The relationship of career mentoring to early career outcomes. *Organization Studies*, 14: 419–441.
- *Whitely, W. T., Dougherty, T. W., & Dreher, G. F. 1991. Relationship of career mentoring and socioeconomic origin to managers' and professionals' early career progress. *Academy of Management Journal*, 34: 331–351.
- Whitener, E. M. 1990. Confusion of confidence intervals and credibility intervals in meta-analysis. *The Journal of Applied Psychology*, 75: 315–321.
- Williams, L. J., & Anderson, S. E. 1991. Job satisfaction and organizational commitment as predictors of organizational citizenship and in-role behaviors. *Journal of Management*, 17: 601–617.
- *Witt, L. A., Kacmar, K. M., & Andrews, M. C. 2001. The interactive effects of procedural justice and exchange ideology on supervisor-rated commitment. *Journal of Organizational Behavior*, 22: 505–515.
- *Wu, P. C., Foo, M. D., & Turban, D. B. 2008. The role of personality in relationship closeness, developer assistance, and career success. *Journal of Vocational Behavior*, 73: 440–448.
- *Yao, X., & Wang, L. 2008. Socially oriented values and reciprocity norm predict organizational commitment. *Asian Journal of Social Psychology*, 11: 247–252.

- *Zatzick, C. D., Elvira, M. M., & Cohen, L. E. 2003. When is more better? The effects of racial composition on voluntary turnover. *Organization Science*, 14: 483–496.
- *Zhang, X., & Bartol, K. M. 2010. The influence of creative process engagement on employee creative performance and overall job performance: A curvilinear assessment. *The Journal of Applied Psychology*, 95: 862–873.
- *Zhang, Z., Waldman, D. A., & Wang, Z. 2012. A multi-level investigation of leader–member exchange, informal leader emergence, and individual and team performance. *Personnel Psychology*, 65: 49–78.



Aparna Joshi (aparnajo@psu.edu) is a professor of management at Smeal College of Business, Pennsylvania State University. She received her PhD from the School of Management and Labor Relations, Rutgers University.

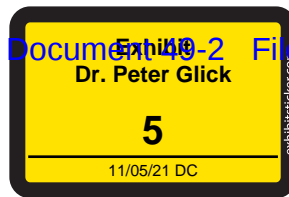
She studies various topics in organizational behavior, including gender, diversity, global dispersion in teams, and generational differences in organizations.

Hyuntak Roh (hroh@yonsei.ac.kr) is an associate professor of management at Yonsei University, Seoul, Korea. He received his PhD from University of Illinois, Urbana-Champaign. His research interests include diversity and inequality issues in the workplace, diversity management practices, and cross-cultural/functional team management.

Jooyeon Son (son22@illinois.edu) received her PhD from the School of Labor and Employment Relations at University of Illinois, Urbana-Champaign, in 2015 and is currently a visiting lecturer there. She conducts research in the areas of gender, team diversity, leader–member exchange, and talent management.



Copyright of Academy of Management Journal is the property of Academy of Management and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.



Are Leader Stereotypes Masculine? A Meta-Analysis of Three Research Paradigms

Anne M. Koenig
University of San Diego

Alice H. Eagly
Northwestern University

Abigail A. Mitchell
Nebraska Wesleyan University

Tiina Ristikari
University of Tampere

This meta-analysis examined the extent to which stereotypes of leaders are culturally masculine. The primary studies fit into 1 of 3 paradigms: (a) In Schein's (1973) *think manager–think male paradigm*, 40 studies with 51 effect sizes compared the similarity of male and leader stereotypes and the similarity of female and leader stereotypes; (b) in Powell and Butterfield's (1979) *agency–communion paradigm*, 22 studies with 47 effect sizes compared stereotypes of leaders' agency and communion; and (c) in Shinar's (1975) *masculinity–femininity paradigm*, 7 studies with 101 effect sizes represented stereotypes of leadership-related occupations on a single masculinity–femininity dimension. Analyses implemented appropriate random and mixed effects models. All 3 paradigms demonstrated overall masculinity of leader stereotypes: (a) In the think manager–think male paradigm, intraclass correlation = .25 for the women–leaders similarity and intraclass correlation = .62 for the men–leaders similarity; (b) in the agency–communion paradigm, $g = 1.55$, indicating greater agency than communion; and (c) in the masculinity–femininity paradigm, $g = 0.92$, indicating greater masculinity than the androgynous scale midpoint. Subgroup and meta-regression analyses indicated that this masculine construal of leadership has decreased over time and was greater for male than female research participants. In addition, stereotypes portrayed leaders as less masculine in educational organizations than in other domains and in moderate- than in high-status leader roles. This article considers the relation of these findings to Eagly and Karau's (2002) role congruity theory, which proposed contextual influences on the incongruity between stereotypes of women and leaders. The implications for prejudice against women leaders are also considered.

Keywords: leadership, management, gender stereotypes, gender roles, meta-analysis

The characteristics that people commonly ascribe to women, men, and leaders contribute to the challenges that women face in obtaining leadership roles and performing well in them.¹ Cultural stereotypes can make it seem that women do not have what it takes for important leadership roles. This cultural mismatch, or *role incongruity*, between women and the perceived demands of leadership underlies biased evaluations of women as leaders (Eagly & Karau, 2002). Fueling this mismatch is an inconsistency between the predominantly communal qualities (e.g., nice, compassionate) that people associate with women and the predominantly agentic

qualities (e.g., assertive, competitive) that they believe are required for success as a leader (Eagly & Carli, 2007). Given that agentic qualities are ascribed more to men than women (e.g., Gallup News Service, 2001; Spence & Buckner, 2000), leadership is generally associated with masculinity. This article reports a meta-analysis of research that has assessed the cultural masculinity of leader stereotypes and explores the conditions under which this masculinity is more or less pronounced.

The Importance of Stereotypes to Women's Access to Leadership

Stereotypes often are a potent barrier to women's advancement to positions of leadership. This assertion is the consensus view not

This article was published Online First May 30, 2011.

Anne M. Koenig, Department of Psychological Sciences, University of San Diego; Alice H. Eagly, Department of Psychology, Northwestern University; Abigail A. Mitchell, Department of Psychology, Nebraska Wesleyan University; Tiina Ristikari, Department of Social Research, University of Tampere, Tampere, Finland.

Anne M. Koenig was supported in part by the National Institutes of Health under a Ruth L. Kirschstein National Research Service Award (1 F31 MH074251).

Correspondence concerning this article should be addressed to Anne M. Koenig, Department of Psychological Sciences, 5998 Alcalá Park, University of San Diego, San Diego, CA 92110. E-mail: akoenig@sandiego.edu

¹ In this article, we use the terms *leader* and *manager* interchangeably. Although a distinction between leadership and management is useful in some contexts (e.g., Bennis, 1989), the research that we review does not allow leader and managerial roles or functions to be separated. Also, we use the terms *sex* and *sexes* to denote the grouping of people into female and male categories. The term *gender* refers to the meanings that societies and individuals ascribe to these female and male categories. We do not intend to use these terms to give priority to any class of causes that may underlie sex and gender effects (see Wood & Eagly, 2010).

only of social and organizational psychologists (e.g., Glick & Fiske, 2007; Heilman & Parks-Stamm, 2007; S. K. Johnson, Murphy, Zewdie, & Reichard, 2008), but also of women who have substantial experience as leaders. For example, a survey of 705 women at the vice president level and above in Fortune 1,000 corporations found that 72% agreed or strongly agreed that “stereotypes about women’s roles and abilities” are a barrier to women’s advancement to the highest levels (Wellington, Kropf, & Gerkovich, 2003).

The damaging effects of stereotypes for women as leaders do not stem from beliefs about women that are mainly negative. On the contrary, consistent with the women-are-wonderful effect (Eagly & Mladinic, 1994; Langford & MacKinnon, 2000), women are regarded as the nicer, kinder sex and thus have a cultural stereotype that is in general more positive than that of men. Nevertheless, women often experience workplace discrimination (Heilman & Eagly, 2008). Theorists have resolved this paradox by reasoning that it is not the evaluative content of the stereotype of women but its mismatch with many desirable work roles that underlies biased evaluations in many employment settings (e.g., Eagly & Karau, 2002; Heilman, 2001; Hogue & Lord, 2007; Lyness & Heilman, 2006).

In Eagly and Karau’s (2002) *role congruity model*, the mismatch that produces biased evaluation is between stereotypes of women (e.g., Spence & Buckner, 2000) and stereotypes of leaders (e.g., Epitropaki & Martin, 2004; Lord & Maher, 1993; Offermann, Kennedy, & Wirtz, 1994). In this model, stereotype content is framed in terms of agency and communion, with communion the more important theme in the female stereotype and agency the more important theme in both leader and male stereotypes. Men are therefore seen as more similar to the leader stereotype than women are, producing disadvantage for women. In Heilman’s (1983, 2001) broader lack-of-fit model, to the extent that a workplace role is inconsistent with the attributes ascribed to an individual, she or he would suffer from perceived lack of fit, producing increased expectations of failure and decreased expectations of success. The incongruity between construals of women and leaders is thus one type of lack of fit.

These communal and agentic meanings ascribed to women and men form a constant backdrop to social interaction, coloring the judgments made about people encountered in organizations and other contexts (Wood & Eagly, 2010). In both the role congruity (Eagly & Karau, 2002) and the lack-of-fit models (Heilman, 2001), these beliefs comprise two kinds of expectations or norms: *descriptive beliefs* (or stereotypes), which are consensual expectations about what members of a social group actually do, and *injunctive* (or *prescriptive*) *beliefs*, which are consensual expectations about what group members ought to do or ideally would do (Cialdini & Trost, 1998). Therefore, women are the targets of two forms of prejudice against them as leaders: a deficit in the ascription of leadership ability to them and, compared to that of men, a less favorable evaluation of their agentic leadership behavior. In other words, descriptively, women seem less usual or natural in most leadership roles; and prescriptively, women often seem inappropriate or presumptuous when they display the agentic behavior often required by these roles (see also Burgess & Borgida, 1999).

Because individuals are commonly assimilated to group stereotypes (e.g., von Hippel, Sekaquaptewa, & Vargas, 1995), an eval-

uative penalty is exacted from women leaders or potential leaders regardless of whether they possess the qualities appropriate for leadership roles. This penalty consists of unfavorable performance expectations, which in turn enable biased judgments and less favorable evaluations (Eagly, Makhijani, & Klonsky, 1992; Lyness & Heilman, 2006). Discrimination is the behavioral outcome of these processes (Eagly & Diekmann, 2005).

How strong is the evidence for the masculinity of leader stereotypes? Partial reviews of research on this question have substantiated the claim about leaders’ cultural masculinity (e.g., Duehr & Bono, 2006; Heilman, Block, Martell, & Simon, 1989; Powell, Butterfield, & Parent, 2002; Schein, 2001). However, these reviews have not identified all of the relevant research paradigms or accessed all of the available studies within each paradigm or estimated the magnitude of the effects. Our integration of the research literature remedies these omissions and tests the propositions of role congruity theory concerning the conditions under which incongruity between the leader and female gender stereotype is stronger or weaker (Eagly & Karau, 2002).

Three Paradigms for Examining the Masculinity of Leader Roles

Research in three separate paradigms has addressed the cultural masculinity of leader stereotypes. Best known is the think manager–think male paradigm, which was created by Schein (1973). This method provides a direct test of the similarity of leader stereotypes to male and female stereotypes. In these studies, separate groups of participants rated a leader category (e.g., successful middle managers), women, or men on a large number of gender-stereotypical traits. The researchers correlated the mean ratings of managers or leaders with the mean ratings of men and, separately, with the mean ratings of women. These correlations represent the similarity of stereotypes of men and women to cultural concepts of leadership. The think manager–think male effect occurs when men and leaders are similar and women and leaders are not similar. Although these studies typically provide only these correlations and not information on the content of the stereotypes of men, women, or leaders, the method yields a direct test of the central assumptions of Eagly and Karau’s (2002) role congruity model and Heilman’s (1983, 2001) lack-of-fit model as applied to leader roles.

A second method, the agency–communion paradigm, consists of studies that examined the gender-stereotypical content of the leader stereotype (Powell & Butterfield, 1979). In these studies, participants rated a leader category (e.g., good manager) on separate masculine (i.e., agentic) and feminine (i.e., communal) gender stereotyping scales. Comparison of participants’ mean ratings on the two scales determined whether the stereotype of leaders was more masculine than feminine. By directly addressing the content of leader roles, this method complements but differs from the similarity comparisons yielded by the think manager–think male paradigm.

A third method, the masculinity–femininity paradigm, consists of studies that appeared in the research literature as a test of the masculine versus feminine content of occupational stereotypes (Shinar, 1975). A substantial minority of the occupations chosen for these studies were leader roles (e.g., university president, mayor). Participants rated each of these leader roles, among other

occupational groups, on a single bipolar masculine versus feminine rating scale. This paradigm became less popular subsequent to critiques pointing out that bipolar masculinity–femininity scales do not allow masculinity and femininity to vary independently but force them to function as opposites (e.g., Constantinople, 1973). Nonetheless, these data are informative in part because researchers investigated various specific types of leaders rather than merely leaders (or managers) in general. If the mean ratings of most categories of leaders were on the masculine side of a masculinity–femininity scale, the studies would provide a conceptual replication of the basic tendency for leader roles to be stereotyped as masculine.

Our project separately meta-analyzed studies in these three paradigms because their distinctively different study designs and measuring instruments precluded combination across the paradigms (see Borenstein, Hedges, Higgins, & Rothstein, 2009, pp. 359–363). The think manager–think male studies correlate (a) ratings of men with ratings of managers and (b) ratings of women with ratings of managers. In this think manager–think male paradigm, participants are randomly assigned to rate one of the three groups (women, men, or leaders). In contrast, in the much simpler agency–communion and masculinity–femininity paradigms, studies obtain ratings only of leaders in general or specific leader roles (and not women or men) on gender-stereotypical traits or scales. Although the agency–communion and masculinity–femininity studies are similar in presenting only a leader group (or groups) to participants, these two paradigms feature different types of measuring instruments. The masculinity–femininity studies use a single bipolar masculine versus feminine rating scale, whereas the agency–communion studies use two separate scales, one assessing agency (or masculinity) and the other assessing communion (or femininity). Therefore, in the masculinity–femininity paradigm, the effect sizes are based on a point estimate (in relation to the scale midpoint). In the agency–communion paradigm, the effect sizes are based on a comparison between ratings on the two scales. For these reasons, the data are noncomparable across the three paradigms and analyzed separately. Yet all three types of studies test the correspondence between gender and leader stereotypes, and the studies are methodologically quite homogenous within each paradigm.

Variation in Stereotypes About Men, Women, and Leaders

Eagly and Karau (2002) hypothesized that the incongruity between leader stereotypes and the female gender stereotype is not fixed but varies with change in either stereotype. This meta-analysis examines several factors hypothesized to influence this incongruity.

Change in Stereotypes Over Time

Cultural change over historical time is one of these factors, given that stereotypes may have changed in a manner that reduces women's role incongruity in relation to leadership. Although some hints of the decreasing masculinity of leadership have appeared in prior reviews (e.g., Duehr & Bono, 2006; Eagly & Sczesny, 2009; Powell et al., 2002; Schein, 2001), generalizations have remained uncertain. Temporal comparisons of relevant studies based on their

dates of publication, which extend back to 1973, allow our meta-analysis to clarify whether the cultural representation of leadership has changed. A shift in an androgynous direction would ease women's role incongruity problem in relation to leader roles.

Why might role incongruity have lessened? Organizational experts have often argued that definitions of good managerial practices have changed in response to features of the contemporary organizational environment, such as fast social and technological change and unprecedented complexity of organizations' missions and contexts (e.g., Avolio, 1999; Kanter, 1997; Lipman-Blumen, 2000). According to such analyses, these changed conditions compromise the efficacy of top-down command-and-control leadership and foster democratic relationships, participatory decision-making, delegation, and team-based leadership skills (e.g., Gergen, 2005; Kanter, 1997; Lipman-Blumen, 2000; McCauley, 2004). Such descriptions are manifestly less masculine than many traditional models of good leadership.

Another possibility is that the increase of women leaders might produce less masculine and more androgynous beliefs about leadership. Evidence of the increase of women leaders abounds, including growth over time in women's emergence as leaders in field and laboratory studies of leader emergence in initially leaderless groups (Eagly & Karau, 1991). In the United States, women now constitute 25% of chief executives when all organizations are considered and 43% of all employees in management, business, and financial operations occupations (vs. 31% in 1983; U.S. Department of Labor, 2007, 2010b). Women have also become more numerous in highly visible political leader roles, now constituting 17% of the U.S. Congress (vs. 2% in 1950; Center for American Women and Politics, 2011) and 12% of governors (vs. 0% in 1950; Center for American Women and Politics, 2010). Women have increased in leadership roles in many other nations as well (see European Commission, 2010; Inter-Parliamentary Union, 2010).

Evidence that the mere presence of more women leaders can change perceptions of leader roles emerged in research on women's occupancy of the chief village councilor role in West Bengal (Beaman, Chattopadhyay, Duflo, Pande, & Topalova, 2009). People in the villages that were mandated (vs. not mandated) by the government to elect a woman for this leader role not only perceived their current women leaders and women leaders in general as more effective but also reduced their implicit bias toward associating men with leadership and elected more women to leadership positions in the next election. As additional evidence that the presence of women leaders changes perceptions of leadership, female college students with more women professors as role models reduced their implicit associations of leadership qualities with men and communal qualities with women (Dasgupta & Asgari, 2004). Empirical research thus has demonstrated that an increase in the number of women leaders can produce a more androgynous concept of leadership and thereby reduce bias toward current and potential women leaders.

What about change in gender stereotypes? If gender stereotypes reflect the differing placements of men and women into social roles (Wood & Eagly, 2010), women's increase in labor force participation (to 61% vs. 33% in 1950; U.S. Department of Labor, 2010a) and in leader roles might predict change in the female stereotype. However, women still perform the majority of domestic work (e.g., Bianchi, Robinson, & Milkie, 2006), and the majority of employed women have remained concentrated in tradi-

tional occupations. The six most common in the United States are secretary and administrative assistant; registered nurse; elementary and middle school teacher; cashier; retail salespersons; and nursing, psychiatric, and home health aides (U.S. Department of Labor, 2011). Therefore, it is not surprising that the partial reviews of gender stereotyping conducted so far have not yielded evidence of decreased stereotyping over time (e.g., Lueptow, Garovich-Szabo, & Lueptow, 2001). Nonetheless, the masculinity of leader roles could be changing without much change in stereotypes about men and women in general.

In sum, leader stereotypes may have become less masculine over time. If the change in leader stereotypes is related to increases of women in leadership roles, then the number of women leaders should be related to the masculinity of leadership. Because stereotype change presumably reflects the updating of impressions based on new observations (e.g., Weber & Crocker, 1983), leader stereotypes at any one time point should correspond best to contemporaneous observations of women in leader roles.

Other Influences on the Masculinity of Leader Roles

A priori moderators. As Eagly and Karau (2002) argued, men may have a more masculine leader stereotype than women do. There is some existing evidence that men, more than women, believe that good leaders have masculine qualities (e.g., Atwater, Brett, Waldman, DiMare, & Hayden, 2004; Schein, 2001) and manifest prejudice against female leaders (Eagly et al., 1992). Such effects are understandable, given that men are less likely than women to have experience with female managers (McTague, Stainback, & Tomaskovic-Devey, 2009; Stainback & Tomaskovic-Devey, 2009) and that men's group interest favors retaining these roles for men. As a result, women leaders, particularly as newcomers entering male-dominated roles, can encounter resistance, especially from men (Eagly & Carli, 2007). Our meta-analysis offers quantitative tests of the extent to which the incongruity between women and leaders is greater in men than women.

Role incongruity should also reflect organizational contexts and the level of leader roles in organizational hierarchies (Eagly & Karau, 2002). Leadership may be less masculine, for example, in female-dominated fields such as elementary education, nursing, or librarianship (U.S. Department of Labor, 2010b). Because these fields are thought to require traditionally feminine skills, such as warmth, compassion, and caring for others (Cejka & Eagly, 1999; Glick, 1991), the characteristics that people associate with leadership roles in these occupations are likely to incorporate more communal attributes. Also, consistent with the preponderance of men in executive positions in many organizations (e.g., 84% of corporate officers in Fortune 500 companies; Catalyst, 2010), construals of leadership are likely to be more masculine for higher status leader positions, thereby increasing role incongruity for women. Our inclusion of studies examining various types of leadership positions allows comparisons of leader stereotypes between different organizational contexts and levels of leadership within organizations.

Comparisons of leader stereotypes across nations are also important, as Eagly and Karau (2002) also argued. Narrative reviewers have suggested that leader roles are less masculine in the United States than in other nations (e.g., Schein, 2001), but there has been no quantitative analysis across nations. Variation in the

percentages of women (vs. men) in leader roles could underlie any national differences as well as temporal differences. The participation of women in leader roles and the overall status of women are considerably greater in Western than Eastern nations (Hausmann, Tyson, & Zahidi, 2009). Thus, there is reason to believe that role incongruity for women leaders is greater in Eastern than Western nations.

In summary, consistent with role congruity theory, we predicted that this meta-analysis would show a robust tendency for leader roles to be perceived as masculine. However, based on this theory and the other evidence outlined above, we expected that the relationship between perceptions of leadership and masculinity would be moderated by several factors. Masculinity of leadership should decrease over time. In addition, men should have a more masculine construal of leadership than women. Also, leadership may be more masculine in domains with few women managers and higher status leader roles. People in different nations may also diverge in their construals of leadership, with the presence of fewer women in leader roles associated with more masculine stereotypes. These a priori moderators of the gender typing of leadership are examined within this meta-analysis.

Exploratory Moderators. Several other variables were also investigated as potential moderators of the masculinity of leadership on an exploratory basis. For example, because employment generally entails some experience with women managers, older participants might have a less masculine concept of leadership. However, older people's more traditional attitudes about gender (e.g., Howell & Day, 2000) might foster a more masculine concept of leadership. Therefore, we could not predict whether age would be positively or negatively associated with the masculinity of leadership but address the issue with an exploratory analysis.

Also, in research on gender, authors of one sex have occasionally obtained different findings than authors of the other sex (e.g., Eagly & Carli, 1981; Leaper, Anderson, & Sanders, 1998). In addition, originators of research paradigms sometimes obtained stronger data than other researchers (e.g., B. T. Johnson & Eagly, 1989; Wood, Lundgren, Ouellette, Busceme, & Blackstone, 1994). Because originators' stronger findings might partially be due to their choice of stereotype measure in the case of the masculinity of leadership, we also explored whether the differing measures used in each paradigm were associated with different effects. Although these variables have influenced other meta-analyses, we had no specific reason to expect they would also influence the masculinity of leadership, so we tested their effects as exploratory moderators.

Method

Sample of Studies and Criteria for Inclusion and Exclusion

Three paradigms of research. The search located studies that fit into the three different paradigms, whose defining characteristics are the following:

1. In the Schein (1973) think manager–think male paradigm, participants rated a leader category, men (typically “men in general”), or women (typically “women in general”) on various traits in a between-subjects design (with only two studies having a within-subjects design). In most of these studies, the ratings were completed using the Schein Descriptive Index, which consists of

92 traits, including many agentic or communal personality traits (e.g., adventurous, submissive, aggressive, intuitive, ambitious, modest, kind). Studies collected for this meta-analysis reported internal consistency coefficients ranging from .71 to .92 for the Schein Descriptive Index. Researchers reported intraclass correlations, computed across the traits, for relating the mean ratings of leaders and men and of leaders and women.² Studies were accepted even if they reported only a men–leader or women–leader correlation, but the great majority of studies reported both correlations.

2. In the Powell and Butterfield (1979) agency–communion paradigm, participants rated leaders (or managers) on two multi-item scales, typically the masculine and feminine scales of the Bem Sex Role Inventory (Bem, 1974). One scale contains stereotypically masculine (i.e., agentic) personality traits, such as assertive, forceful, dominant, and competitive; the other scale contains stereotypically feminine (i.e., communal) personality traits, such as affectionate, compassionate, warm, and gentle. Studies collected for this meta-analysis reported internal consistency coefficients for the Bem Sex Role Inventory from .74 to .90 for the agency scale and .79 to .90 for the communion scale. Researchers usually reported means and standard deviations on the two scales but sometimes presented item-level data that included a representation of agentic and communal traits or classified respondents' leader ratings into the four quadrants of a two-dimensional agency–communion space. The researchers other than Powell and Butterfield whose studies fit this paradigm usually did not cite the precedent of the Powell and Butterfield study, but nonetheless reported participants' ratings of a leader category on agency and communion.

3. In the Shinar (1975) masculinity–femininity paradigm, participants rated one or more leader categories on a single bipolar 7-point scale that ranged from *very masculine* to *very feminine*. Researchers reported means and standard deviations of these ratings and compared means to the midscale value, which represents the gender-neutral or androgynous point on the scale. The meta-analysis includes only those occupations from each study that have a clear leadership focus (most commonly managerial roles, such as office manager).

The studies selected for all three of these paradigms presented participants with a general leader category, such as managers or executives, or with occupations or job titles denoting managerial authority, such as personnel directors, head librarians, academic administrators, military officers, or political office holders. For all paradigms, data were excluded if they provided ratings of leader behaviors rather than personality traits (e.g., Gutek & Stevens, 1979), of specific people in a leadership position (e.g., Petty & Miles, 1976), or of leader groups identified by sex (e.g., “successful female managers”; Dodge, Gilroy, & Fenzel, 1995). Studies in the agency–communion paradigm were eliminated if the majority of the items in their two scales did not pertain to agentic or communal personality traits or if the two types of items were very unequally represented (e.g., Buttner & Rosen, 1988).

The abstract of each identified document was evaluated by at least one of the authors; if the study potentially fit the inclusion criteria, the full document was obtained. In total, 78 documents were rejected based on their abstracts and 134 documents were rejected after reading the document. In addition, 11 documents met the inclusion criteria but did not report the appropriate statistics to

calculate an effect size, and when contacted, the authors could not supply the needed information (e.g., Couch & Sigler, 2001; Gerstner & Day, 1994).

Studies from any participant population or nation were included if they fit the above criteria. When documents included data from different nations or different participant samples within a nation, they were treated as separate studies if the results were reported separately. Some documents included data for more than one type of leader category (e.g., company president, head librarian). If separate groups of participants rated each category, the leader categories were treated as separate studies. If the same participants rated more than one leader category, these effect sizes were averaged prior to calculating study-level mean effect sizes but retained as separate effect sizes for moderator analyses. In two documents in the think manager–think male paradigm (Dodge et al., 1995; Karau, 2005), separate groups of participants rated different types of leaders. These ratings were correlated with the same ratings of women and men, provided by two other groups of participants. In our analyses, we treated these semi-independent leader conditions as separate studies.

Some data qualified in more than one paradigm. For example, we extracted only the leader condition of think manager–think male studies for use in the agency–communion paradigm if the report contained item-level data that allowed us to separate agentic and communal items into subscales (i.e., van Engen, 2006). Also, parts of the same study that were published separately sometimes qualified for different paradigms. For example, some authors presented intraclass correlation coefficients in one publication but grouped the data into agentic and communal scales in another (Fullagar, Sverke, Sverke, Sümer, & Slick, 2003; Sümer, 2006) or presented the intraclass correlation coefficients in one publication but the mean for a specific masculine–feminine bipolar scale in another (Koch, Luft, & Kruse, 2005; Luft, 2003).

Search for studies. Computer-based information searches were conducted in the following databases: ABI/INFORM, Academic Press/Ideal, Business Source Elite, Proquest Digital Dissertations, Educational Resources Information Center (ERIC), Emerald Full Text, PAIS International, Proquest Business Databases, PsycINFO, ScienceDirect, Sociological Abstracts, Web of Science, WilsonWeb, and Worldwide Political Science Abstracts.

In these searches, the keywords *leader** (-s, -ship), *manager** (-s, -ial), *educator*, *executive*, *candidate*, *public office*, *political office*, *principal*, or *occupation* were combined with (a) *stereotyp** (-e, -es, -ical), *traits*, *characterization*, *attribute inventory*, *image*,

² The design of the studies in this paradigm appears similar to that of a study by Broverman, Broverman, Clarkson, Rosenkrantz, and Vogel (1970), whose participants rated a “mature, healthy, socially competent” man, woman, or adult person. Broverman et al. found greater similarity between an adult person and men than between an adult person and women, but this finding has been criticized as artifactual (see Kelley & Blashfield, 2008; Widiger & Settle, 1987, for details). However, the similarity of the Broverman et al. study to the think manager–think male studies is only superficial because Schein’s (1973) method offers superior item selection and statistical analysis. Specifically, (a) the typical think manager–think male items (in the Schein Descriptive Index) are moderately balanced between agentic and communal qualities (see Duehr & Bono, 2006), and (b) the statistical analysis with an intraclass correlation is more appropriate to the data.

expectation, and *perce** (-*ption*, -*tive*) or (b) gender terms, such as *man*, *men*, *woman*, *women*, *male*, *female*, *masculinity*, *femininity*, *sex*, *sex role*, *sex-typing*, *gender*, *gender role*, and *androg** (-*ny*, -*nous*). Additional searches paired the two most common dependent measures (the Schein Descriptive Index and the Bem Sex Role Inventory) with *manage** or *leader**, *think manager–think male*, or *requisite management characteristics*. Web of Science citation searches were also performed for the seminal articles in each paradigm (Powell & Butterfield, 1979; Schein, 1973; Shinar, 1975). All obtained studies' reference sections were also searched for relevant studies.

Several foreign psychological and academic databases were searched with the limited keywords *leader** (-*s*, -*ship*) or *manager** (-*s*, -*ial*) paired with (a) *stereotyp** (-*e*, -*es*, -*ical*), *traits*, or *similarities* or (b) gender terms, such as *masculinity*, *femininity*, *gender*, and *sex* (translated, as appropriate). The foreign databases were from Finland (University of Joensuu database, Forum of Science database, Finnish Social Science Achieve, University of Rovaniemi database, University of Tampere database); Germany (Datenbank, PSYINDEX, PSYINDEXalert, PSYTKOM); Great Britain and Ireland (Index to Theses); Norway (National Library of Norway, BIBSYS Library); Spain (Psicodoc); and Sweden (Chalmers University of Technology database, the Center for Research Libraries). Documents in languages other than English were translated by appropriately skilled individuals who assisted the two coders.

To locate unpublished studies, messages asking for relevant data were sent to the e-mail distribution lists of several organizations: European Association of Social Psychology, European Association of Work and Organizational Psychology, Interamerican Society of Psychology, Society for the Psychology of Women, Social Issues in Management Division of the Academy of Management, Society for Personality and Social Psychology, and Society for the Psychological Study of Social Issues. Data were also sought from the originators of each of the paradigms, but they had no additional data to offer. Finally, when the authors of this article gave several talks on the meta-analysis in progress, they asked the audience to contribute their own data or to provide leads about relevant sources of data. Both unpublished and published studies meeting the inclusion criteria were included in the meta-analysis.

The initial search extended from the beginning of each paradigm through the end of 2002. The search was then updated in June 2007 by consulting the databases that had yielded studies in 2002, with the addition of Google and Google Scholar, and updated once more in March 2009 with PsycINFO, Google, and Google Scholar.

Variables Representing Each Study in All Paradigms

The coded variables represent the *a priori* moderators as well as other variables that we investigated on an exploratory basis. The included studies, their codings, and effect sizes appear in Tables 1, 2, and 3 for the think manager–think male, agency–communion, and masculinity–femininity paradigms, respectively.

A priori moderators. Year of publication (or year of data collection for unpublished data) was recorded as well as the percentage of male participants. Effect sizes were also calculated separately by participant sex when possible. If data were not reported separately for male and female participants, we contacted the authors and requested this information.

For each paradigm, we recorded the exact description of the leader group and coded its domain: (a) In the think manager–think male paradigm, the leader domains were nearly all designated as managerial, and those few not so designated were heterogeneous (e.g., leader, educational administrator), so no moderator variable was tested; (b) in the agency–communion paradigm, the domains were managerial (e.g., manager, middle-level manager), political (e.g., mayor, state senator), or educational (e.g., elementary school principal, school superintendent); and (c) in the masculinity–femininity paradigm, the domains, or economic sectors, of the occupations were business (e.g., company president, marketing manager), education (e.g., educational administrator, school principal), politics (mayor, politician), judicial (e.g., federal judge, Supreme Court justice), arts (orchestra conductor, theatrical director), or other (boat captain, farm manager, park manager). For all three paradigms, leader status was coded as high (defined as prestigious political roles such as president or governor or organizational positions higher than middle manager such as upper-level managers or executives) or moderate (e.g., managers, middle-level managers, all other leader occupations).

Participant nationality was coded by the nation where the study was conducted and classified as Western (United States, Great Britain, Germany, Canada, New Zealand, Australia, Sweden, South Africa, Netherlands, Denmark, and Portugal) or Eastern (China, Japan, Hong Kong, Egypt, Saudi Arabia, and Turkey). Other exploratory classifications (e.g., North American vs. other Western vs. Eastern) did not improve prediction of the effect sizes.

To interpret cross-temporal and cross-national comparisons of studies, we imported data on the percentage of female managers, defined as the percentage of managers who are women, based on the likely time of data collection for the nations in which the studies were conducted, defined as 2 years prior to studies' dates of publication. When available, these data came from the *Human Development Reports* of the United Nations Development Programme (e.g., 2006) and otherwise from the *Yearbooks of Labour Statistics* of the International Labour Organization (e.g., 1985). Aggregate indexes of the status of women available from the United Nations Development Programme or other sources were not appropriate for this meta-analysis because they are available only for more recent years.

Exploratory moderators. Average participant age was either provided in each research report or estimated from sample characteristics. The distributions of the effect sizes across the types of participant populations were as follows: (a) for the think manager–think male paradigm, 55% undergraduate students, 18% managers, 8% MBA students, 6% nonmanagerial employees, and 14% other or mixed; (b) for the agency–communion paradigm, 62% undergraduate students, 13% nonmanagerial employees, 9% managers, 9% MBA students, and 9% other or mixed; and (c) for the masculinity–femininity paradigm, 86% undergraduate students and 14% other or mixed.

The percentage of male authors was coded. Research group was coded as originators of paradigm (Schein, Powell and Butterfield, or Shinar) or other. Stereotype measure was coded based on the use of the originators' versus other measures: (a) In the think manager–think male paradigm, Schein Descriptive Index or other (e.g., task vs. person-orientation scales); (b) in the agency–communion paradigm, Bem Sex Role Inventory or other (e.g., initiating structure vs. consideration scales); and (c) in the masculinity–femininity paradigm, always a masculinity–femininity 7-point scale.

Table 1
All Included Studies (k = 51) With Moderator Values and Effect Sizes for the Think Manager–Think Male Paradigm

Report	Sample ^a	Sep. by sex ^b	Pub. source ^c	% male authors	Stereotype measure ^d	Nation ^e	% female mgr.	M age	Leader role ^f	Leader status ^g	n	% male part.	Item n	Leader sim. ^h	ICC ⁱ	Variance
Booyens & Nkomo (2006)	1	Y	1	0	1	ZA	30	32	SMM	1	391	69	92	W	.46	0.01
Boyce & Herd (2003)	1	Y	1	0	1	US	46	20	SMO	1	404	69	92	M	.70	0.01
Brenner et al. (1989)	1	Y	1	67	1	US	35	42	SMM	1	501	82	92	M	-.03	0.01
Byler (2000)	1	Y	2	0	1	US	44	44	EAD	1	407	82	92	M	.39	0.01
Cohen-Kaner (1992)	1	Y	2	0	1	US	37	36	SMM	1	395	69	92	M	.15	0.01
Dion & Schuller (1990)	1	Y	1	50	2	US/CA	36	32	SMM	1	83	82	92	W	.68	0.01
Dion & Schuller (1991)	1	Y	1	50	2	US/CA	37	33	SMM	1	85	78	92	W	.45	0.01
Dodge et al. (1995)	1	Y	1	33	1	US	40	31	SMM	1	119	39	92	M	.23	0.01
Duehr & Bono (2003)	1	Y	2	0	1	US	42	30	SMM	1	119	39	92	W	.26	0.01
Duehr & Bono (2006)	2	Y	1	0	1	US	42	47	SMM	1	119	39	92	M	.53	0.01
Fernandes & Cabral-Cardoso (2003)	1	Y	1	50	2	PT	32	21	M	1	53	45	51	M	.58	0.02
Foster (1994)	1	Y	1	0	1	GB	32	42	SMM	1	43	48	22	M	.66	0.05
Fullager et al. (2003)	2	Y	1	50	1	TR	8	21	SMM	1	48	59	92	W	.15	0.01
Heilman et al. (1989)	1	M	1	50	1	SE	30	21	SMM	1	48	59	92	M	.50	0.25
Karau (2005)	1	Y	2	100	1	US	42	22	SMM	1	48	59	92	W	.02	0.01
Karau & Elsaid (2005)	2	Y	2	100	1	US	42	22	SMM	1	48	59	92	M	.59	0.01
Karau & Hansen (2005)	2	Y	2	100	1	SE	31	21	SMM	1	76	39	92	W	.54	0.01
Kent (1984)	1	Y	2	0	1	US	28	47	SSA	1	74	31	92	M	.56	0.01
											178	60	92	W	.67	0.01
											169	57	92	M	.56	0.01
											197	30	92	W	.28	0.01
											203	33	92	M	.43	0.01
											199	46	42	W	-.01	0.03
											199	46	42	M	.40	0.03
											160	50	92	W	.11	0.01
											160	50	92	M	.46	0.01
											536	58	92	W	.21	0.01
											536	58	92	M	.58	0.01
											182	27	92	W	.15	0.01
											182	27	92	M	.68	0.01
											78	100	92	W	-.24	0.01
											78	100	92	M	.54	0.01
											112	59	92	W	.01	0.01
											104	63	92	M	.39	0.01
											113	58	92	W	.19	0.01
											105	63	92	M	.93	0.01
											217	58	92	W	.31	0.01
											220	59	92	M	.40	0.01
											371	75	92	W	-.24	0.01
											357	73	92	M	.43	0.01
											47	55	92	W	.44	0.01
											47	57	92	M	.33	0.01
											61	51	92	W	.28	0.01
											61	56	92	M	.39	0.01
											240	47	92	W	.58	0.01
											247	52	92	M	.59	0.01

MASCULINITY OF LEADER STEREOTYPES

623

Table 1 (continued)

Report	Sample ^a	Sep. by sex ^b	Pub. source ^c	% male authors	Stereotype measure ^d	Nation ^e	% female mgr.	M age	Leader role ^f	Leader status ^g	n	% male part.	Item n	Leader sim. ^h	ICC ⁱ	Variance
Kunkel et al. (2003)	1	Y	1	33	1	US	46	21	CEO	2	41	46	92	W	.18	0.01
Luft (2003)	1	Y	2	0	2	DE	27	28	L	1	42	46	92	M	.80	0.01
Marmenout (2008)	1	Y	2	0	1	SA	31	21	SMM	1	202	47	25	W	-.27	0.04
Massengill & DiMarco (1979)	1	Y	1	100	1	US	22	44	SM	1	144	40	92	W	.68	0.04
Mook (2005)	1	Y	2	0	1	NL	26	39	SM	1	171	37	92	M	.18	0.01
Neergaard et al. (2007)	1	Y	2	33	1	DK	25	43	MSC	1	89	60	92	W	.61	0.01
Norris & Wylie (1995)	1	Y	1	0	2	US/CA	40	21	SMM	1	83	57	92	W	.15	0.01
Orser (1994) ^j	2	Y	1	0	1	CA	41	21	SM	1	85	48	92	M	.64	0.01
Paradine et al. (1995)	1	Y	2	66	2	CA	41	21	SM	1	96	45	92	W	.24	0.01
Ryan et al. (in press)	2	Y	2	25	1	US	37	22	MSC	1	119	85	92	M	.58	0.01
Sauers et al. (2002)	1	Y	1	67	1	NZ	38	21	SMM	1	113	82	92	W	.44	0.01
Schein (1973)	1	M	1	0	1	US	18	43	SMM	1	613	48	24	W	.72	0.01
Schein (1975)	1	F	1	0	1	US	17	44	SMM	1	605	48	24	M	.81	0.05
Schein & Davidson (1993) ^k	2	Y	1	0	1	GB	32	21	SMM	1	198	0	92	W	.31	0.01
Schein & Mueller (1992)	2	Y	1	50	1	GB	32	21	SMM	1	198	0	92	M	.46	0.01
Schein et al. (1989)	1	Y	1	33	1	US	35	21	SMM	1	198	50	92	W	.18	0.01
Schein et al. (1996)	2	Y	1	50	1	JP	9	21	SMM	1	198	50	92	M	.49	0.01
Sczesny (2003a)	1	Y	1	0	2	DE	27	24	L	1	204	55	40	W	.15	0.03
											213	54	40	M	.76	0.03
											35	38	92	W	.36	0.01
											36	51	92	M	.72	0.01
											75	25	92	W	.46	0.01
											212	48	92	M	.79	0.01
											212	48	92	W	.43	0.01
											193	100	92	W	.71	0.01
											212	100	92	M	.06	0.01
											113	0	92	W	.62	0.01
											115	0	92	M	.30	0.01
											152	54	92	W	.54	0.01
											152	54	92	M	.19	0.01
											101	48	92	W	.52	0.01
											101	48	92	M	.64	0.01
											101	48	92	W	.17	0.01
											101	48	92	M	.64	0.01
											297	63	92	W	.05	0.01
											297	63	92	M	.71	0.01
											152	64	92	W	.23	0.01
											152	64	92	M	.64	0.01
											195	66	92	W	-.06	0.01
											210	64	92	M	.60	0.01
											186	56	92	W	.10	0.01
											183	53	92	M	.91	0.01
											73	58	40	W	.65	0.03
											73	58	40	M	.76	0.03

(table continues)

Table 1 (continued)

Report	Sample ^a	Sep. by sex ^b	Pub. source ^c	% male authors	Stereotype measure ^d	Nation ^e	% female mgr.	M age	Leader role ^f	Leader status ^g	n	% male part.	Item n	Leader sim. ^h	ICC ⁱ	Variance
Szeszsy (2003b)	1	Y	1	0	2	DE	27	45	L	1	60	58	40	W	.87	0.03
Sylvan (1983)	1	Y	2	0	1	US	27	38	SM	1	60	53	92	W	.87	0.03
van Engen (2006)	1	Y	2	0	1	NL	26	22	SMM	1	70	23	92	W	.11	0.01
Yim & Bond (2002)	1	Y	1	50	2	HK	25	19	SMM	1	31	0	92	W	.48	0.01
											142	49	32	W	.29	0.01
											142	49	32	M	.38	0.01
											142	49	32	M	.44	0.03
											142	49	32	M	.72	0.03

Note. In each sample, separate groups of participants rated a leader category, women, or men on gender-stereotypical traits. The effect sizes are the intraclass correlations between the mean ratings of managers (or leaders) and (a) the mean ratings of men or (b) the mean ratings of women across all the traits. Sep. = separated; pub. = publication; mgr. = managers; part. = participants; sim. = similarity; ICC = intraclass correlation coefficient.

^a The number of different samples (e.g., from different populations or countries) reported within the article. ^b Separated samples by sex: N = no; Y = yes; M = sample was all male; F = sample was all female. ^c Publication source: 1 = published; 2 = unpublished. ^d 1 = Schein Descriptive Index; 2 = other. ^e AU = Australia; CA = Canada; CN = China; DK = Denmark; EG = Egypt; DE = Germany; GB = Great Britain; HK = Hong Kong; JP = Japan; NL = Netherlands; NZ = New Zealand; PT = Portugal; SA = Saudi Arabia; ZA = South Africa; SE = Sweden; TR = Turkey; US = United States; X = included data from several countries. ^f EAD = effective athletic directors; L = leadership; M = managers; MSC = managers of successful companies; CEO = successful CEOs; SM = successful managers; SMM = successful middle managers; SMO = successful military officers; SSA = successful school administrators; SUM = successful upper level managers. ^g 1 = moderate; 2 = high. ^h M = men-leaders similarity; W = women-leaders similarity. ⁱ Effect size (ICC) computed using a one-way single rater random effects model, related the mean ratings of the leader role to the ratings of men or women. These correlations represent the similarity of stereotypes of men or women to stereotypes of leaders; higher correlations indicate a greater similarity. ^j The sample with 50% male participants was composed of business students, and the sample with 0% male participants was composed of nonbusiness students. ^k The sample with 54% male participants was from North Great Britain, and the sample with 48% male participants was from South Great Britain.

Lastly, the source of publication was coded as published or unpublished (including dissertations, master's theses, and other unpublished documents) to provide one estimate of possible publication bias (Sutton, 2009; see the subsection "Publication Bias").

Two of the authors, who had successfully completed a course on meta-analysis, independently coded the studies, with high agreement across the variables. The mean kappas for intercoder agreement were .90, .90, and .97 for the think manager-think male, agency-communion, and masculinity-femininity paradigms, respectively. Disagreements were resolved by discussion.

Computation of Effect Sizes and Data Analysis

Effect sizes were calculated with a hand calculator or DSTAT software and then entered into Comprehensive Meta-Analysis (Version 2.2.050) and Statistical Package for the Social Sciences (SPSS). Subgroup analyses were conducted using Comprehensive Meta-Analysis, and meta-regressions were conducted with SPSS macros provided by D. Wilson (see <http://mason.gmu.edu/~dwilsonb/ma.html>). For all analyses, *p* values less than .05 were considered significant and values between .05 and .10 were defined as marginal.

Calculation of effect sizes. In the think manager-think male paradigm, researchers reported intraclass correlation coefficients (ICCs) for the men-leaders and women-leaders relationships. In all cases in which ICCs were not reported, the researchers provided additional information that allowed us to produce intraclass correlations. The ICCs used in the primary data reports were computed using a one-way, single-rater, random effects model, which assesses absolute agreement among measurements (Case 1 in McGraw & Wong, 1996). However, the more appropriate ICC would have been the 2-way ICC(A,1), which is also a measure of absolute agreement but which takes the fixed column factor (leaders vs. women; leaders vs. men) into account. Thus, the ICC provided by researchers in the primary studies would be biased downward somewhat but quite close to the calculations for ICC(A,1) (K. O. McGraw, personal communication, January 8, 2003). We used the one-way, single-rater, random effects ICC because it was available for all data sets, providing a common metric for the studies. The men-leaders and women-leaders similarities, or ICCs, were analyzed separately as Fisher's *Z*, using the conversion $.5 * \log\left(\frac{1 + [(k - 1) * r]}{1 - r}\right)$, where *k* is the number of observations made on each object of measurement, and given an inverse variance within-study weight of $1/[k/2 * (n - 2)(k - 1)]$, where *n* is the number of items (see formulas in Appendix B of McGraw & Wong, 1996). The *Z*s were transformed back to ICCs for presentation of the results.³ In the random-effects models, the variances in these study weighting terms consisted of the sum of the within-study variance and the between-studies variance (see Borenstein et al., 2009, p. 72).

In the agency-communion paradigm, researchers reported means and standard deviations separately on the agentic and communal scales, allowing the computation of a *d* effect size comparing the

³ Because the women-leaders and men-leaders correlations are dependent in that they share the data for leaders, we did not compare the two effects statistically. These types of comparisons would require information about the correlation between male and female targets, which is not available in the primary studies.

MASCULINITY OF LEADER STEREOTYPES

625

Table 2
All Included Studies (k = 48) With Moderator Values and Effect Sizes for the Agency-Communion Paradigm

Report/leader role	Sample ^a	Sep. by sex ^b	Pub. source ^c	% male authors	Stereotype measure ^d	Nation ^e	% female mgr.	M age	Leader domain ^f	Leader status ^g	n	% male part.	g^h	Variance
Butterfield & Powell (1981) Good president	1	N	1	100	1	US	25	20	2	2	128	59	2.36	0.03
Cann & Siegfried (1987) Manager	1	Y	1	100	2	US	33	21	3	1	110	53	0.79	0.02
Huddy & Terkildsen (1993) Good Congress member	1	Y	1	0	2	US	37	21	2	2	133	54	0.48	0.02
Good local council member									2	1	149	50	0.20	0.01
Good mayor									2	2	153	51	0.75	0.01
Good president									2	2	139	56	1.00	0.02
Average of roles													0.61	0.02
Inderlied & Powell (1979) Good manager	4	Y	1	50	1	US	22	43	3	1	41	63	2.65	0.09
Good manager								37	3	1	51	0	2.14	0.06
Good manager								28	3	1	154	69	1.41	0.02
Good manager								22	3	1	259	58	1.22	0.01
Maier (1993) Typical manager	1	Y	1	100	1	US	37	20	3	1	60	50	0.72	0.04
Average of roles									3	1	59	50	1.79	0.05
Ideal manager													1.25	0.04
Marongiu & Ekehammar (2000) Manager	1	Y	1	50	1	SE	27	41	3	1	87	48	0.65	0.02
Powell (1992) Good manager	2	Y	1	100	1	US	37	21	3	1	87	47	1.46	0.03
Typical American manager									3	1	79	44	1.63	0.03
Powell & Butterfield (1979) Good manager	2	Y	1	100	1	US	22	20	3	1	110	82	2.35	0.20
Good manager								27	3	1	574	70	1.98	0.05
Powell & Butterfield (1984) Good manager	1	N	1	100	1	US	28	20	3	1	627	62	2.47	0.01
Powell & Butterfield (1987) Good president of the U.S.	2	Y	1	100	1	US	33	20	2	2	50	46	1.53	0.05
Good vice-president									2	2	44	46	0.81	0.05
Powell & Butterfield (1989) Good manager	2	Y	1	100	1	US	35	20	3	1	199	43	2.56	0.15
Good manager								27	3	1	126	60	1.97	0.23
Powell et al. (2002) Good manager	2	Y	1	67	1	US	45	32	3	1	123	65	1.49	0.13
Good manager								21	3	1	206	57	1.08	0.04
Powell & Kido (1994) Good manager	2	Y	1	10	1	JP	8	21	3	1	88	77	-1.18	0.03
Typical Japanese manager								21	3	1	86	23	0.27	0.02
Rosenwasser & Dean (1989) City council member	9	N	1	0	1	US	35	21	2	1	14	54	2.44	0.25
Governor									2	2	14	54	2.66	0.27
Mayor									2	2	14	54	2.31	0.24

(table continues)

Table 2 (continued)

Report/leader role	Sample ^a	Sep. by sex ^b	Pub. source ^c	% male authors	Stereotype measure ^d	Nation ^e	% female mgr.	M age	Leader domain ^f	Leader status ^g	n	% male part.	g ^h	Variance
President														
School board member													1.40	0.18
State representative													1.01	0.16
State senator													2.40	0.25
U.S. representative													1.59	0.19
U.S. senator													2.55	0.26
Rustemeyer & Thrien (1989)	2	Y	1	0	1	DE	19						1.98	0.21
Good manager								23	3	1	109	64	3.11	0.25
Good manager								45	3	1	54	100	1.84	0.23
Stoker (2007)	1	Y	1	0	1	NL	26							
Ideal manager														
Stümer (2006)	1	Y	1	0	2	TR	8							
Successful middle manager								36	3	1	3229	67	1.07	0.001
van Engen (2006)	1	Y	2	0	2	NL	26							
Successful middle manager								21	3	1	369	55	1.29	0.01
Vinnicombe & Cames (1998)	1	Y	1	0	2	LU	9							
Successful manager at their bank								43	3	1	66	50	1.02	0.03
Vinnicombe & Singh (2002)	1	Y	1	0	2	GB	33							
Successful managers who reached the top team								43	3	2	363	33	1.71	0.01
Willemssen (2002)	1	Y	1	0	2	NL	27							
Successful manager at their bank								22	3	1	139	52	1.19	0.02
Williams (1989)	1	N	2	0	1	US	35							
Academic dean of students								50	1	1	54	43	1.00	0.04
Assistant principal									1	1	54	43	0.76	0.04
Elementary school principal									1	1	54	43	0.28	0.04
School superintendent									1	2	54	43	0.34	0.04
Secondary school principal									1	1	54	43	1.20	0.04
Average of roles													0.71	0.04

Note. In each sample, participants rated a leader role on separate masculine (i.e., agentic) and feminine (i.e., communal) gender stereotyping scales. Comparison of participants' mean ratings on the two scales determined whether the stereotype of leaders was more masculine than feminine. The effect size for the average of the roles is presented for studies using a within-subjects design. Sep. = separated; pub. = publication; mgr. = managers; part. = participants.

^a The number of different samples (e.g., from different populations or countries) reported within the article. Some studies with multiple samples reported participant characteristics across all samples. ^b Separated samples by sex: N = no; Y = yes. ^c Publication source: 1 = published; 2 = unpublished. ^d 1 = Bem Sex Role Inventory; 2 = other. ^e DE = Germany; GB = Great Britain; JP = Japan; LU = Luxembourg; NL = Netherlands; SE = Sweden; TR = Turkey; US = United States. ^f 1 = educational; 2 = political; 3 = managerial. ^g 1 = moderate; 2 = high. ^h Effect size (*g*) computed by subtracting the mean rating on the feminine scale from the mean rating on the masculine scale and dividing by the pooled standard deviation, adjusted for small sample bias. If only frequencies or percentages in a 2 × 2 table created by splitting the scales at the medians were reported, agency and communion were treated as dichotomous, and *g* was estimated from *d*_{cor}. Larger *g*s indicate a more masculine concept of leadership.

MASCULINITY OF LEADER STEREOTYPES

627

Table 3
All Included Studies (k = 101) With Moderator Values and Effect Sizes for the Masculinity-Femininity Paradigm

Report/leader role	Sample ^a	Sep. by sex ^b	% male authors	Nation ^c	% female mgr.	M age	Leader domain ^d	Leader status ^e	n	% male part.	g ^f	Variance
Beggs & Doolittle (1993)	1	Y	0	US	37	21	6	1	139	50	2.44	0.007
Boat captain							2	2	138	49	1.22	0.007
Company president							1	1	140	50	0.61	0.007
Dean/educational administrator							2	1	141	50	0.88	0.007
Director, labor relations							2	1	141	50	1.11	0.007
Director, operations							6	1	140	50	1.64	0.007
Farm manager							5	1	141	50	1.46	0.007
Federal judge							1	1	141	50	-1.66	0.007
Head librarian							2	1	141	50	0.87	0.007
Manager, credit union							2	1	141	50	0.72	0.007
Manager, procurement services							6	1	139	50	0.85	0.007
Managing editor							4	1	141	50	1.40	0.007
Mayor							2	1	141	50	0.45	0.007
Motel manager							2	1	141	50	0.26	0.007
Office manager							3	1	140	50	1.35	0.007
Orchestra conductor							6	1	141	50	1.10	0.007
Park manager							2	1	139	50	0.30	0.007
Personnel director							4	1	140	51	1.40	0.007
Politician							2	1	141	50	0.06	0.007
Public relations director							2	1	140	50	0.60	0.007
Sales manager							2	2	140	50	1.35	0.007
Sales president							1	1	141	50	0.92	0.007
School principal							3	1	141	50	0.01	0.007
Theatrical director							1	2	140	50	1.44	0.007
University president							5	1	139	50	1.51	0.007
U.S. Supreme Court justice										0.89	0.007	0.007
Average of roles												
Gatton et al. (1999)	1	Y	33	US	44	21						
Accounting manager							2	1	175	47	0.84	0.006
Computer operations manager							2	1	175	47	0.67	0.006
Data processing manager							2	1	175	47	-0.16	0.006
Engineering manager							2	1	175	47	1.46	0.006
Financial manager							2	1	175	47	0.76	0.006
Human resources manager							2	1	175	47	-0.31	0.006
Marketing manager							2	1	175	47	0.20	0.006
Office manager							2	1	175	47	0.27	0.006
Plant manager							2	1	175	47	1.77	0.006
Quality control manager							2	1	175	47	0.84	0.006
R & D director							2	1	175	47	0.45	0.006
Sales manager							2	1	175	47	0.38	0.006
Average of roles										0.60	0.006	0.006
Koch et al. (2005)	1	Y	0	DE	27	28						
Leadership							2	1	202	47	0.63	0.005
											(table continues)	

Table 3 (continued)

Report/leader role	Sample ^a	Sep. by sex ^b	% male authors	Nation ^c	% female mgr.	M age	Leader domain ^d	Leader status ^e	n	% male part.	g ^f	Variance
Mahoney & Blake (1987)	1	Y	100	US	33	26						
Farm manager							6	1	29	52	1.52	0.036
Financial manager							2	1	34	50	0.48	0.030
Sales manager							2	1	29	52	0.58	0.035
Average of roles											0.86	0.033
Muñoz Sastre et al. (2000)	2	N	20	FR	9	21						
Boat captain							6	1	234	45	2.35	0.004
Company president							2	2	234	45	0.98	0.004
Dean/educational administrator							1	1	234	45	0.92	0.004
Farm manager							6	1	234	45	1.97	0.004
Federal judge							5	1	234	45	0.53	0.004
Head librarian							1	1	234	45	-1.03	0.004
Managing editor							2	1	234	45	0.52	0.004
Mayor							4	1	234	45	0.78	0.004
Motel manager							2	1	234	45	0.80	0.004
Orchestra conductor							3	1	234	45	1.22	0.004
Park manager							6	1	234	45	0.93	0.004
Personnel director							2	1	234	45	0.49	0.004
Politician							4	1	234	45	1.15	0.004
Public relations director							2	1	234	45	0.28	0.004
Sales manager							2	1	234	45	0.30	0.004
Sales president							2	2	234	45	0.72	0.004
School principal							1	1	234	45	0.07	0.004
Supreme Court justice							5	1	234	45	0.60	0.004
Theatrical director							3	1	234	45	0.53	0.004
University president							1	2	234	45	0.87	0.004
Average of roles											0.75	0.004
				ES	32	21						
Boat captain							6	1	126	41	2.03	0.008
Company president							2	2	126	41	0.75	0.008
Dean/educational administrator							1	1	126	41	0.74	0.008
Farm manager							6	1	126	41	1.25	0.008
Federal judge							5	1	126	41	0.59	0.008
Head librarian							1	1	126	41	-0.09	0.008
Managing editor							2	1	126	41	0.74	0.008
Mayor							4	1	126	41	0.61	0.008
Motel manager							2	1	126	41	0.68	0.008
Orchestra conductor							3	1	126	41	1.30	0.008
Park manager							6	1	126	41	0.78	0.008
Personnel director							2	1	126	41	0.70	0.008
Politician							4	1	126	41	0.62	0.008
Public relations director							2	1	126	41	0.07	0.008
Sales manager							2	1	126	41	0.51	0.008
Sales president							2	2	126	41	0.45	0.008
School principal							1	1	126	41	0.40	0.008
Supreme Court justice							5	1	126	41	0.88	0.008

MASCULINITY OF LEADER STEREOTYPES

629

Table 3 (continued)

Report/leader role	Sample ^a	Sep. by sex ^b	% male authors	Nation ^c	% female mgr.	M age	Leader domain ^d	Leader status ^e	n	% male part.	g ^f	Variance
Theatrical director							3	1	126	41	0.32	0.008
University president							1	2	126	41	0.57	0.008
Average of roles											0.70	0.008
Shinar (1975)	1	N	0	US	18	21						
Boat captain							6	1	120	50	3.91	0.009
Company president							2	1	120	50	9.06	0.011
Dean/educational administrator							1	1	120	50	0.99	0.008
Farm manager							6	1	120	50	2.09	0.008
Federal judge							5	1	120	50	1.76	0.008
Head librarian							1	1	120	50	-1.59	0.008
Managing editor							2	1	120	50	0.68	0.008
Mayor							4	1	120	50	1.73	0.008
Motel manager							2	1	120	50	1.08	0.008
Orchestra conductor							3	1	120	50	2.73	0.009
Park manager							6	1	120	50	3.03	0.009
Personnel director							2	1	120	50	0.50	0.008
Politician							4	1	120	50	0.96	0.008
Public relations director							2	1	120	50	0.68	0.008
Sales manager							2	1	120	50	1.13	0.008
Sales president							2	2	120	50	1.62	0.008
School principal							1	1	120	50	0.67	0.008
Theatrical director							3	1	120	50	1.09	0.008
University president							1	2	120	50	3.94	0.009
U.S. Supreme Court justice							5	1	120	50	7.08	0.010
Average of roles											2.16	0.009

Note. In each sample, participants rated leader roles on a single bipolar masculinity–femininity scale. Comparison of these ratings to the midpoint of the scale determined whether the stereotype of leaders was masculine or feminine. The effect size for the average of the roles is presented for studies using a within-subjects design. All studies were published and used a 1–7 masculine–feminine rating scale. Sep. = separated; mgr. = managers; part. = participants.

^a The number of different samples (e.g., from different populations or countries) reported within the article. ^b Separated samples by sex: N = no; Y = yes. ^c FR = France; DE = Germany; ES = Spain; US = United States. ^d 1 = education; 2 = business; 3 = arts; 4 = political; 5 = judicial; 6 = other. ^e 1 = moderate; 2 = high. ^f Effect size (g) computed by subtracting the midpoint of the scale from the leader rating, divided by the standard deviation and corrected for small sample bias.

ratings on the two scales: $(M_1 - M_2)/s_p$. The effect sizes were converted to g with the correction for small sample bias: $1 - [3/(4N - 9)]$ (Borenstein et al., 2009). Some authors split their sample at the median on both scales and reported the frequencies or percentages in each quadrant of the resulting 2×2 table. If only this report was available, agency and communion were treated as dichotomous, and g was estimated from d_{Cox} , which is a logistic transformation of the odds-ratio (Sánchez-Meca, Marín-Martínez, & Chancón-Moscoso, 2003, Formula 18). The within-study weighting term was the conventional inverse variance for standardized comparisons of means (Lipsey & Wilson, 2001, p. 72) or d_{Cox} (Sánchez-Meca et al., 2003, Formula 19), with the random-effects models also incorporating the between-studies variances in the study weight.

In the masculinity–femininity paradigm, authors reported means and standard deviations on a 7-point masculinity–femininity scale, which yielded a d as a comparison with the scale's midpoint $[(M - \text{midpoint})/SD]$. The effect sizes were then corrected for small sample bias with the formula $1 - [3/(4N - 5)]$ to create a g (Borenstein et al., 2009). The within-study weighting term for these effect sizes was $(1/n) + \{(d * d)/[2n(n - 1)]\}$ (B. Becker, personal communication, June 19, 2008), and the random-effects models included the between-studies variance in the study weight.

Mean and distribution of effect sizes. Within each paradigm, the presentation first considers the mean weighted effect size (with within-subjects effect sizes combined prior to averaging), calculated by a random-effects model because the studies' effect sizes were not assumed to be consistent with a single underlying mean value. We tested for outliers, as defined by more than 1.5 times the interquartile range beyond the upper quartile (i.e., the upper inner fence, Tukey, 1977). We used Cohen's (1988) benchmarks for d and r to describe the magnitude of the effect sizes g and ICC, under which a g of .20 or an ICC of .10 is considered small, a g of .50 or an ICC of .30 is considered medium, and a g of .80 or an ICC of .50 is considered large. Along with the overall mean, we present several distributional statistics recommended by Borenstein et al. (2009): (a) T (or tau, the estimated standard deviation of the true effect sizes); (b) 95% confidence interval (CI; a measure of the accuracy of the mean; 95% of mean effect sizes would fall inside this interval); (c) 90% prediction interval (PI; a measure of the dispersion of effect sizes; 90% of true effects in new studies with a sample selected at random would fall inside this interval); (d) Q (a measure of uncertainty, or whether heterogeneity is genuine); and (e) I^2 (a measure of the magnitude of heterogeneity, defined as the proportion of the observed dispersion that is real rather than spurious on a 1 to 100% scale). An I^2 near zero indicates that almost all of the observed variance in the effects is spurious and that there is no variance to explain, whereas a large I^2 indicates that investigation of the reasons for this variance is warranted.

Publication bias. Analyses of the distribution of effect sizes checked for potential biases in publication or our retrieval of studies (see Borenstein et al., 2009). We first examined the funnel plot of the effect sizes plotted by the standard error and assessed whether Egger's test of the plot's asymmetry was significant. We next implemented the trim-and-fill procedure, which estimates the number of studies that should be removed to create a more symmetric funnel plot. This procedure assesses the impact of removing these studies on the mean effect size, then fills these studies back in and imputes a mirror image for each such study to correct the

variance. In addition, a small studies analysis using cumulative meta-analysis was used to calculate the mean effect as each study was added, starting with those with the largest sample sizes. This analysis determines whether the mean effect size estimate has stabilized with the larger studies and does not shift when smaller studies are added (see Borenstein et al., 2009). Finally, a mixed-effects subgroup analysis tested whether published and unpublished studies differed in their mean effect size.

Moderator analyses. We present subgroup analyses and meta-regressions to examine whether moderators accounted for variability among effect sizes (using each effect size as a separate sample). For each subgroup analysis, we present the overall mean and tau for each subgroup. Within each paradigm, we first examined the effect sizes as moderated by the theory-relevant, a priori variables (i.e., publication year, participant sex, leader domain, leader status, participant nationality, and percentage of female managers) and then as moderated by the other, exploratory variables. In each paradigm, all coded variables were tested as single predictors. However, although the results for all a priori variables are presented for each paradigm, only exploratory variables that produced significant moderation as single predictors are reported.

The subgroup analyses used mixed-effects models because these subgroups involve specific grouping (such as males and females) that are fixed across studies (Borenstein et al., 2009; Lipsey & Wilson, 2001). These mixed models consisted of a random-effects model within subgroups and a fixed-effect model across subgroups, with appropriate inverse variance weights (Borenstein et al., 2009). Meta-regressions were calculated as random-effects models, with appropriate inverse variance weights (Borenstein et al., 2009). These calculations were simple meta-regressions for each centered continuous moderator (e.g., year of publication) and a multiple meta-regression for the moderators that were individually significant in the subgroup analyses or simple meta-regressions. The categorical moderators were recoded into two groups for entry into each multiple meta-regression based on maximizing the differences between the subgroup categories, as determined by the single-variable moderator analyses. The random- and mixed-effects models were calculated using the method of moments (Borenstein et al., 2009).

Results

Think Manager–Think Male Paradigm

Mean and distribution of effect sizes. Table 1 lists the 40 studies with their characteristics and 51 effect sizes. A more masculine stereotype is indicated by (a) a greater men–leaders similarity, as indexed by a larger men–leaders ICC; and (b) a lesser women–leaders similarity, as indexed by a smaller women–leaders ICC. The weighted mean effect for the 51 men–leaders ICCs was .62 ($T = 0.26$; 95% CI [.57, .66]; 90% PI [.27, .82]). There was a large amount of variability within the effect sizes, $Q(50) = 321.65$, $p < .001$; $I^2 = 84.46$. The weighted mean effect for the 49 women–leaders ICCs was .25 ($T = 0.22$; 95% CI [.19, .32]; 90% PI [−.12, .56]). There was also a large amount of variability within these effect sizes, $Q(48) = 241.77$, $p < .001$; $I^2 = 80.15$. Thus, consistent with an overall think manager–think male effect, the men–leaders similarity was large and the women–leaders similarity was small. Although the variability in each of these two types

of effect sizes was large, as indicated by T and the confidence and prediction intervals, a high proportion of the variability reflected real differences, as indicated by I^2 .⁴

Publication bias. Analyses checked for potential biases in publication or our retrieval of studies for both the women–leaders similarity and the men–leaders similarity. Egger’s test of the plot’s asymmetry was not significant for either of the two sets of effect sizes, suggesting no asymmetry. For the men–leaders similarity, under the assumption that biases would hinder the publication of studies with small effects, the trim-and-fill procedure suggested that no studies be trimmed. For the women–leaders similarity, under the assumption that biases would hinder the publication of studies with large effects (because they would be contrary to the hypothesis of low similarity between women and leaders), the trim-and-fill procedure suggested that four studies with small effect sizes should be trimmed to make the plot more symmetric, increasing the mean effect size from .25 to .29. Computing these analyses with the effect size weights based on the number of participants rather than the number of items also revealed no significant asymmetry in the funnel plots, and no studies were trimmed. In addition, using either the number of items or the number of participants to order the data, cumulative meta-analyses from large to small studies showed very little movement of the effect size with the addition of small studies. Finally, a mixed-effects subgroup analysis comparing published and unpublished studies yielded no difference for the women–leaders similarity, $Q_B(1) = 0.28$, $p = .60$, or the men–leaders similarity, $Q_B(1) = 1.00$, $p = .32$. Thus, overall, there appears to be little publication bias, and the bias that may exist would only slightly increase the women–leaders similarity.

Analyses for each moderator. These analyses report the impact of each moderator on the women–leaders similarity and the men–leaders similarity, as shown in Table 4 for the subgroup analyses and Table 5 for the simple meta-regressions. Models are presented for both women–leaders and men–leaders effect sizes for each theory-relevant variable, but they are presented for other exploratory variables only when the model yielded a p value of .05 or less for either type of similarity.

Publication year (ranging from 1973 to 2010) was marginally related to the women–leaders similarity, with earlier years associated with a weaker women–leaders similarity, but was not significantly related to the men–leaders similarity. Male (vs. female) participants were associated with a weaker women–leaders similarity, but this variable was not significantly related to the men–leaders similarity. With participant samples not divided by sex, greater percentages of male participants were associated with a marginally weaker women–leaders similarity but not significantly related to the men–leaders similarity.⁵

High-status (vs. moderate-status) leaders were associated with a stronger men–leaders similarity but not significantly related to the women–leaders similarity. Eastern (vs. Western) participant nationalities were associated with a marginally weaker women–leaders similarity but not significantly related to the men–leaders similarity. Smaller percentages of female managers in participants’ nations were not significantly related to the women–leaders or men–leaders similarity.

Among the exploratory moderators, older participant age was associated with a stronger women–leaders similarity but not significantly related to the men–leaders similarity. Greater percent-

ages of male authors were associated with a weaker women–leaders similarity but not significantly related to the men–leaders similarity.

Multiple meta-regression analyses. Given some confounding between the predictors, it is important to examine each moderator controlled for the effects of the other moderators. Table 6 shows models that regressed both sets of effect sizes onto the predictors that were at least marginally significant in any of the subgroup or simple meta-regression models. However, this analysis omitted participant sex because it was created by splitting studies’ participant samples. In the resulting model, the women–leaders similarity was stronger in later publication years, with a lower percentage of male participants, and with older participants. The men–leaders similarity was stronger for higher status leaders.

Agency–Communion Paradigm

Mean and distribution of effect sizes. Table 2 lists the 22 studies with their characteristics and 48 effect sizes. Larger g s indicate a more masculine stereotype, as defined by the agency minus communion difference. One sample (Powell & Kido, 1994; ratings of “good manager”) was removed from further analyses as an outlier ($g = -1.18$). The weighted mean effect for the 39 remaining samples (merging the data within each of the three studies that used within-subjects designs) was $g = 1.55$ ($T = 0.55$; 95% CI [1.36, 1.75]; 90% PI [0.61, 2.49]). Thus, the overall effect was strongly in the masculine direction, with participants rating leader groups (e.g., managers) as higher in agentic than communal traits, with the lower bounds of the confidence and prediction intervals remaining above zero. Although there was a large amount of variability in the effect sizes, $Q(38) = 615.79$, $p < .001$, I^2 was 93.83, showing that a high proportion of the variability reflected real differences, not random error.⁶

Publication bias. The funnel plot of the effect sizes plotted by the standard error was asymmetric by Egger’s test of asymmetry, with smaller studies clustering to the right (masculine direction) of the mean effect size. The trim-and-fill procedure suggested that 14 studies with large effect sizes should be trimmed to make the plot more symmetric, decreasing the mean effect size from 1.55 to 1.15. A cumulative meta-analysis from large to small studies indicated very little change in the overall effect when the smallest studies

⁴ We also recomputed the analyses using the number of participants as the weight with generally similar results. The overall effects and associated statistics were nearly identical, although the Q and I^2 were larger for both the women–leaders and men–leaders similarities. Subgroup analyses and meta-regressions produced findings similar to those obtained with item n used to compute the weights.

⁵ Schein (2001; Schein, Mueller, & Jacobson, 1989) maintained that the think manager–think male effect has decreased over time only for female participants. However, although the women–leaders similarity was weaker for female than male participants and increased over time, there was no evidence of greater increase for female than male participants.

⁶ The weighted mean effect including the outlier was 1.49 ($T = 0.64$; 95% CI [1.27, 1.71]; 90% PI [0.40, 2.58]). $Q(39) = 835.88$, $p < .001$; $I^2 = 95.33$. We also computed the same moderator analyses including the outlier. The effects were similar to those reported here, although the leader domain model was only marginal and the models for percentage of male participants, percentage of male authors, and research group were nonsignificant.

Table 4
Subgroup Analyses for the Think Manager–Think Male Paradigm

Variable and class	Women–leaders similarity					Men–leaders similarity				
	Q_B	p	k	ICC	T	Q_B	p	k	ICC	T
Participant sex	30.38	<.001				1.43	.23			
Male			47	.11	0.20			48	.63	0.26
Female			47	.37	0.24			49	.58	0.27
Leader status	0.82	.37				8.72	.003			
Moderate			46	.26	0.23			48	.60	0.22
High			3	.13	0.00			3	.82	0.48
Participant nationality	2.88	.09				0.95	.33			
Western			42	.27	0.22			44	.60	0.24
Eastern			6	.09	0.18			6	.68	0.38

Note. Q_B = between-classes effect with degrees of freedom equal to the number of moderator levels minus one; p = probability; k = number of samples; ICC = mean weighted effect size represented by the intraclass correlation coefficient; T = tau, estimated standard deviation of the true effect sizes.

were included. Finally, a mixed-effects subgroup analysis comparing published and unpublished studies (albeit limited by the small number of unpublished studies) indicated that the mean effect was marginally larger for published studies ($g = 1.59$, $T = 0.56$, $k = 37$) than unpublished studies ($g = 0.86$, $T = 0.00$, $k = 2$), $Q_B = 2.91$, $p = .09$. Thus, the impact of publication bias was not trivial; nevertheless, the finding that leader stereotypes were more agentic than communal remains valid.

Analyses for each moderator. Table 7 presents the subgroup analyses, and Table 8 presents the simple meta-regression models (left columns). Moderation by participant nationality was not tested because there were only two non-Western samples. Among the a priori moderators, publication year (ranging from 1979 to 2007) was significant, with earlier years associated with larger agency–communion differences. Greater percentages of male participants were associated with larger differences, although dividing the sample by participant sex (leaving only 25 of 47 reports) did

not yield a significant model. Leader domains that were managerial or political were associated with larger differences than the educational domain, but leader status did not significantly moderate the effect. Percentage of female managers also did not significantly moderate the difference. Among the exploratory moderators, the originators' (i.e., Powell or Butterfield) research group (vs. others) was associated with larger differences, as was the originators' stereotype measure (vs. others).

Multiple meta-regression analysis. Table 8 (right columns) shows a model that regressed the agency–communion differences onto the variables that were significant predictors in the subgroup or simple meta-regression models. In this model, the agency–communion difference was larger for managerial and political (vs. educational) domains and with the originators' stereotype measure and marginally larger in earlier years.

Masculinity–Femininity Paradigm

Mean and distribution of effect sizes. Table 3 lists the seven studies with their characteristics and 101 effect sizes. Larger g s indicate a more masculine stereotype in relation to the 0.00 that denotes the androgynous midpoint of the scale. If we regard all of the occupations within each study as separate samples, nine samples were outliers. Above the mean (i.e., masculine direction), the outlier occupations were company president, U.S. Supreme Court justice, university president, boat captain, park manager, and orchestra conductor. Below the mean (i.e., feminine direction), the outlier occupations were three instances of head librarian. All of these effect sizes were retained in the analyses because their large distances from the mean were due to the understandable tendency for very male-dominated or female-dominated managerial occupations to yield extreme means and small standard deviations.

The weighted mean effect of the seven studies, produced by first averaging across the ratings of different leadership roles within each study was $g = 0.94$ ($T = 0.51$; 95% CI [0.56, 1.32]; 90% PI [−0.15, 2.03]). Thus, the overall effect is strongly in the masculine direction, with participants rating occupations entailing leadership as quite masculine. There was a large amount of variability in the effect sizes, $Q(6) = 223.09$, $p < .001$. Nevertheless, I^2 was 97.31,

Table 5
Simple Meta-Regressions for the Think Manager–Think Male Paradigm

Predictor	k	b	β	p
Year of publication				
Women–leaders similarity	49	0.01	0.22	.09
Men–leaders similarity	51	0.00	0.04	.77
% Male participants				
Women–leaders similarity	49	−0.00	−0.22	.09
Men–leaders similarity	51	0.00	0.06	.66
% Female managers				
Women–leaders similarity	48	0.00	0.18	.18
Men–leaders similarity	50	−0.00	−0.16	.27
Participant age				
Women–leaders similarity	49	0.01	0.28	.03
Men–leaders similarity	51	−0.00	−0.04	.75
% Male authors				
Women–leaders similarity	49	−0.00	−0.29	.02
Men–leaders similarity	51	0.00	0.08	.56

Note. Models are random-effects weighted linear regressions calculated with weights equal to the reciprocal of the variance for each effect size plus a random-effects component. k = number of samples; p = probability; b = unstandardized regression coefficient; β = standardized regression coefficient.

Table 6
Multiple Meta-Regressions for the Think Manager–Think Male Paradigm

Variable and class	Women–leaders similarity			Men–leaders similarity		
	<i>b</i>	β	<i>p</i>	<i>b</i>	β	<i>p</i>
Year of publication	0.01	0.47	<.001	–0.00	–0.05	.75
% Male participants	–0.00	–0.27	.03	0.00	0.07	.62
Leader status	–0.11	–0.10	.37	0.49	0.43	.002
Participant nationality	–0.11	–0.13	.27	0.17	0.19	.18
Participant age	0.01	0.47	.001	0.00	0.02	.90
% Male authors	–0.00	–0.13	.31	–0.00	–0.02	.91

Note. Models are random-effects weighted linear regressions calculated with weights equal to the reciprocal of the variance for each effect size plus a random-effects component. *b* = unstandardized regression coefficient; β = standardized regression coefficient; *p* = probability. *k* = 48 for women–leaders similarity and *k* = 50 for men–leaders similarity. Codes were as follows: leader status: 1 = high, 0 = moderate; participant nationality: 1 = Eastern, 0 = Western.

showing that a high proportion of the variability reflected real differences, not random error.⁷

Publication bias. It is unclear how publication bias would operate for this paradigm because the participants rated a heterogeneous set of occupations and the researchers did not compare managerial occupations to other types of occupations. Thus, there would be no bias against publication depending on whether managerial occupations were rated as more or less masculine. Nonetheless, we analyzed the distribution of effect sizes for potential biases in publication or study retrieval. With the funnel plot of the effect sizes plotted by the standard error with each occupation represented separately, Egger’s test was significant, but the trim-and-fill procedure did not suggest that any effect sizes be trimmed. When we first averaged the occupations within each study, Egger’s test was not significant, and the trim-and-fill method did not suggest that any studies be trimmed. A cumulative meta-analysis

from large to small studies indicated very little change in the overall effect when the smallest study was included. Published and unpublished studies could not be compared because all studies were published. Overall, there is thus little evidence of publication bias.

Analyses for each moderator. Table 9 presents the subgroup analyses, and Table 10 presents the simple meta-regression models (left columns). Limited sample variability on many study characteristics precluded testing all but a few as potential moderators. For example, there were no non-Western samples, all but one sample had young (i.e., student) participants, and all studies used the same dependent measure.

Among the a priori moderators, publication year (ranging from 1975 to 2005) was significant, with earlier years associated with stronger masculinity. Greater percentages of male participants were also associated with stronger masculinity, although dividing the samples by participant sex (leaving only 41 of 101 effect sizes) did not yield a significant model. In terms of leader domain, other leader roles (boat captain, farm manager, park manager) and judicial roles were associated with stronger masculinity, followed by political, arts, and business roles, and finally by educational roles. High-status (vs. moderate-status) leaders were associated with stronger masculinity. Smaller percentages of female managers in participants’ nations were associated with marginally stronger masculinity. Among the exploratory moderators, smaller percentages of male authors were associated with stronger masculinity.

Table 7
Subgroup Analyses for the Agency–Communion Paradigm

Variable and class	Q_B	<i>p</i>	<i>k</i>	<i>g</i>	<i>T</i>
Participant sex	0.00	.99			
Male			25	1.20	0.33
Female			25	1.20	0.52
Leader domain	7.69	.02			
Educational			6	0.75 _b	0.33
Political			15	1.48 _a	0.74
Managerial			26	1.51 _a	0.54
Leader status	0.43	.51			
Moderate			39	1.43	0.57
High			8	1.27	0.60
Research group	10.13	.001			
Originator			18	1.68	0.64
Other			29	1.17	0.37
Stereotype measure	7.48	.006			
Originator’s			37	1.55	0.64
Other			10	0.95	0.48

Note. Q_B = between-classes effect with degrees of freedom equal to the number of moderator levels minus one; *p* = probability; *k* = number of samples; *d* = mean weighted effect size; *T* = tau, estimated standard deviation of the true effect sizes. Leader domain *ds* not sharing subscripts are significantly different from each other at *p* < .05 when entered into a mixed-effects model.

⁷ Removing the nine outlying effect sizes still resulted in a significantly masculine mean weighted effect size of 0.82 (*T* = 0.19; 95% CI [0.67, 0.97]; 90% PI [0.41, 1.23]), $Q(6)$ = 35.44, *p* < .001; I^2 = 83.07. The moderator analyses were similar to those reported here, although the models for leader status and percentage of female managers were nonsignificant. In addition, we computed the effect sizes based on a mean metric, using the *M* on the 1–7 scale as the effect size. We estimated the standard error of each effect from the standard deviation divided by the square root of the sample size, yielding a weight of one divided by the squared standard error (see Lipsey & Wilson, 2001). Using this metric, the mean-weighted effect size was masculine (*M* = 4.91; *T* = 0.58; 95% CI [4.48, 5.35]; 90% PI [3.65, 6.17]), $Q(6)$ = 4647.70, *p* < .001; I^2 = 99.87. The moderator analyses were also very similar using this metric.

Table 8
Meta-Regressions for the Agency–Communion Paradigm

Predictor	<i>k</i>	Simple meta-regressions			Multiple meta-regression		
		<i>b</i>	β	<i>p</i>	<i>b</i>	β	<i>p</i>
Year of publication	47	−0.04	−0.38	.004	−0.02	−0.26	.08
% Male participants	47	0.01	0.27	.05	0.01	0.11	.41
Leader domain	47				0.91	0.43	.004
% Female managers	47	−0.01	−0.09	.47			
Research group	47				−0.01	−0.01	.97
Stereotype measure	47				0.54	0.32	.03

Note. Models are random-effects weighted linear regressions calculated with weights equal to the reciprocal of the variance for each effect size plus a random-effects component. Blank cells indicate data that are not applicable to that analysis. *k* = number of samples; *b* = unstandardized regression coefficient; β = standardized regression coefficient; *p* = probability. Codings were as follows: leader domain: 1 = managerial and political leaders, 0 = educational leaders; research group: 1 = originator, 0 = other; stereotype measure: 1 = originator's, 0 = other.

Multiple meta-regression analysis. Table 10 (right columns) shows a multiple meta-regression model that regressed the masculinity–femininity effect sizes onto the variables that were at least marginally significant predictors in the subgroup and simple meta-regression models. In this model, leader occupations were more masculine in earlier publication years, in judicial or other domains, and with high-status roles.

Discussion

All three paradigms showed that stereotypes of leaders are decidedly masculine. Specifically, people viewed leaders as quite similar to men but not very similar to women, as more agentic than communal, and as more masculine than feminine. These findings are minimally invalidated by publication bias. Only the agency–communion findings appeared somewhat susceptible to exaggeration of the masculinity of the leader stereotype, but this weak bias

does not compromise the strong overall masculinity of the leader stereotype.

In addition to this large and highly robust masculinity effect, our analyses revealed considerable variability in the extent to which stereotypes of leaders are masculine versus more androgynous, although the rarity of negative effect sizes in the agency–communion and masculinity–femininity paradigms indicates that conceptions of leadership hardly ever strayed into feminine territory. This moderation for the most part was consistent with our predictions. To examine the extent to which the findings supported this theory, we briefly review and discuss them across the three paradigms.

A Priori Moderators

Year of publication. Evidence of the increasing androgyny of the leader stereotype over publication years emerged in all three paradigms, including in the multiple regression equations that controlled for other moderator variables (albeit as a marginal effect in the agency–communion multiple meta-regression). In the think manager–think male paradigm, the increasing similarity of leaders and women that emerged in our analyses could be due to change in women or leader stereotypes or both. Because participants rated only leaders (and not men or women) in the agency–communion and the masculinity–femininity paradigms, the effects of moderators on these effect sizes could be due only to differing beliefs about leaders. Given that all three paradigms showed increasing androgyny of leadership over time and that previous research indicates little change in stereotypes of women toward greater masculinity (see Lueptow et al., 2001), the most appropriate conclusion appears to be that the popular conception of leadership has changed over time. In addition, the think manager–think male paradigm indicated that change over time took the form of increasing similarity between leaders and women without change in the similarity between leaders and men. Thus, our conclusion is that leadership now, more than in the past, appears to incorporate more feminine relational qualities, such as sensitivity, warmth, and understanding, thus adding them to the masculine dominance and strength qualities traditionally associated with leadership.

Table 9
Subgroup Analyses for the Masculinity–Femininity Paradigm

Variable and class	<i>Q_B</i>	<i>p</i>	<i>k</i>	<i>g</i>	<i>T</i>
Participant sex	2.53	.11			
Male			41	0.92	0.73
Female			41	0.67	0.70
Leader domain	18.81	.002			
Education			16	0.49 _a	1.21
Business			48	0.84 _a	1.02
Arts			8	1.07 _{a,b}	0.77
Political			8	1.08 _{a,b}	0.37
Judicial			8	1.80 _{b,c}	1.87
Other ^a			13	1.93 _c	0.85
Leader status	8.36	.004			
Moderate			89	0.92	0.96
High			12	1.91	1.97

Note. *Q_B* = between classes effect with degrees of freedom equal to the number of moderator levels minus one; *p* = probability; *k* = number of samples; *d* = mean weighted effect size; *T* = tau, estimated standard deviation of the true effect sizes. Leader domain *d*s not sharing subscripts are significantly different from each other at *p* < .05 when entered into a mixed-effects model.

^a Other occupations were boat captain, farm manager, park manager.

Table 10
Meta-Regressions for the Masculinity–Femininity Paradigm

Predictor	<i>k</i>	Simple meta-regressions			Multiple meta-regression		
		<i>b</i>	β	<i>p</i>	<i>b</i>	β	<i>p</i>
Year of publication	101	–0.06	–0.42	<.001	–0.05	–0.39	<.001
% Male participants	101	0.08	0.22	.01	–0.01	–0.02	.85
Leader domain	101				1.17	0.36	<.001
Leader status	101				1.20	0.30	<.001
% Female managers	101	–0.02	–0.16	.06	0.00	0.00	.98
% Male authors	101	–0.01	–0.19	.02	–0.00	–0.02	.76

Note. Models are random-effects weighted linear regressions calculated with weights equal to the reciprocal of the variance for each effect size plus a random-effects component. Blank cells indicate data that were not applicable to that analysis. *k* = number of samples; *b* = unstandardized regression coefficient; β = standardized regression coefficient; *p* = probability. Codings were as follows: leader domain: 1 = judicial and other, 0 = education, business, arts, and political; leader status: 1 = high, 0 = moderate.

Consistent with existing research (Beaman et al., 2009; Dasgupta & Asgari, 2004), this change may reflect people's increasing exposure to women among organizational managers and political leaders. Yet this change may also reflect the increasing failure of traditionally masculine, command-and-control styles of leadership to meet the complex challenges involved in contemporary management and political leadership (Kanter, 1997; Lipman-Blumen, 2000). If so, men leaders may be under pressure to enlarge their behavioral repertoire to include a greater measure of culturally feminine relational skills (Eagly & Carli, 2007). Men appear to be succeeding, given that they are not perceived as any less similar to leaders than they were in the past, according to our analysis that separated the role incongruity of men and leaders from that of women and leaders in the think manager–think male paradigm. These findings of increased femininity of leadership without decreased masculinity suggest that women leaders would be well advised to retain elements of a masculine leadership style to avoid a mismatch with leader roles, even if they now have greater flexibility to incorporate elements of a feminine leadership style.

Participant sex. An additional prediction was that men would have a more masculine construal of leadership than women. Only the think manager–think male paradigm allowed all of the findings to be separated or identified by participant sex, and the resulting analyses yielded smaller women–leaders correlations for male than female participants. Although this finding was not replicated in the other paradigms, a related analysis that did not rely on separating female and male participant samples showed that a greater proportion of male participants yielded marginally smaller women–leaders correlations in the think manager–think male paradigm as well as significantly larger masculine leader stereotypes in both the agency–communion and the masculinity–femininity paradigms. In sum, evidence that men construe leadership as more masculine than women is present in all three paradigms and strongest for men's stereotyping of leaders as not very similar to women.

This finding is consistent with meta-analytic research showing that men, but not women, devalued women's leadership in experimental studies that held constant all leader characteristics other than their sex (see Eagly et al., 1992). Also, men's failure to accord women many of the qualities of leaders is consistent with their greater social dominance (e.g., Pratto, Sidanius, & Levin, 2006)

and sexism (e.g., Glick & Fiske, 1996). The implications of men's reservations about women's leadership abilities are important. Because men are the more typical holders of decision-making power in organizations, their beliefs that women lack the qualities of leaders can limit women's access to positions of authority and slow their progress into leadership positions.

Leader status. Consistent with the rarity of women in top positions, higher status leadership positions were expected to have a more masculine stereotype. Both the think manager–think male and the masculinity–femininity paradigms found greater masculinity of high-status leadership positions (e.g., upper-level managers) than moderate-status positions (e.g., middle-level managers or managers in general). The think manager–think male studies suggested that this effect appears mainly because higher status leaders are stereotypically more similar to men rather than less similar to women. Although these effects remained significant in the multiple meta-regressions, they should be interpreted cautiously because the studies offered relatively few instances of especially high-status leader roles. In addition, the effects may be constrained to certain types of high-status roles, as the agency–communion paradigm, which contained a greater number of political high-status leadership positions than the other two paradigms, did not yield this moderation.

According to role congruity theory, women entering higher status positions should encounter more prejudice because of the greater stereotypical mismatch between women and leadership. Our think manager–think male findings suggest instead that men gain greater congruity as status increases, thereby contributing to their increasing numerical dominance at higher levels. Yet women who do manage to be successful in very high-status roles may be perceived as highly competent, because people assume a double standard whereby such women had to overcome especially difficult challenges (Rosette & Tost, 2010).

Other a priori variables. Our ability to test the remaining a priori moderators—leader domain, nationality of participants, and percentage of female managers—was compromised by the limitations of the available studies. Because the three paradigms used different types of leader groups, the effects of leader domain were not comparable across the paradigms, and the think manager–think male paradigm did not present sufficient variability in domain to allow a test of this variable. In the agency–communion and

masculinity–femininity paradigms, educational positions were the least masculine, consistent with the female domination of the education administrator job category (63% women; U.S. Department of Labor, 2010b). In the masculinity–femininity paradigm, the judicial domain and the category of “other” leaders were the most masculine. Judicial occupations (supreme court justice and federal judge) are highly male dominated: Only four women have ever served as Supreme Court Justices, and only 9% of federal judges serving since 1789 have been women, including 30% of currently active sitting judges (U.S. Courts, 2011). However, the “other” leader domain, consisting of boat captain, farm manager, and park manager, deserves comment. These heterogeneous roles are likely also male dominated, as men constitute 82% of farm, ranch, and other agricultural managers and 79% of people in farming, fishing, and forestry occupations in general (U.S. Department of Labor, 2010b). In addition, these are outdoor occupations that often require physical strength as well as leadership abilities, which likely increased their perceived masculinity (Cejka & Eagly, 1999). Future research should systematically investigate stereotypes in leader domains that vary in the representation of women and other attributes.

Comparisons of effects across nations were limited by the overrepresentation of the United States and the rarity of studies from non-Western nations. Nevertheless, consistent with national data on the status of women (e.g., Hausmann et al., 2009) and on attitudes toward women as leaders (P. Norris & Inglehart, 2001), participants from Eastern nations showed a marginally weaker women–leaders similarity than participants from Western nations (although this effect was nonsignificant in the multiple regression). The percentage of female managers in participants’ nations was a more informative variable, which was associated with marginally weaker masculinity in the masculinity–femininity paradigm (although a nonsignificant predictor in the multiple meta-regression). Yet this variable was not significantly related to masculinity in the think manager–think male and agency–communion paradigms. A more systematic global representation of nations would provide a better test of the impact of women in management on leader stereotypes.

Exploratory Moderators

The exploratory variables yielded a few findings of interest. In the think manager–think male paradigm, older participants, who were a mix of managerial and nonmanagerial employees, showed a stronger women–leaders similarity than younger participants, who were business or other undergraduate students. Although these results support the conclusion that students and others with little workplace experience hold more masculine stereotypes of leaders, the nonreplication in the agency–communion paradigm renders these findings ambiguous.

The originators of the agency–communion paradigm produced more masculine stereotypes, although this effect was no longer significant in the multiple regression and not replicated in the other paradigms. The originator’s measure may have produced greater masculinity in the agency–communion paradigm, because this measure, the Bem Sex Role Inventory, had better psychometric properties than many of the other measures, which were often more casually constructed without a factor or item analyses. Yet the rarity of reporting measures’ internal consistencies in the

primary studies precluded a statistical test of the idea that the Bem Sex Role Inventory provided more coherent measures of agency and communion. The percentage of male authors had inconsistent effects across the paradigms and was not a significant predictor the multiple meta-regressions. In general, the lack of consistent moderation of these paradigms by variables related to the authors of the primary studies and to stereotype measures used in the studies indicates that the effects revealed in this meta-analysis are not limited to certain research groups or dependent measures.

Strengths and Limitations of the Paradigms

Our three-part meta-analysis demonstrates the advantages of examining leader stereotypes from the perspective of differing research paradigms. The fact that the paradigms conceptually replicate one another lends confidence to results that are similar across the paradigms. However, each paradigm has its own strengths and limitations.

The think manager–think male paradigm has the advantage of comparing the cultural construal of leadership to that of men and women separately, thus allowing for leadership to be seen as similar to men, women, both, or neither. The ICC also provides a direct measure of the similarity between female and male stereotypes and leader stereotypes, yielding a clear-cut test of role congruity theory (Eagly & Karau, 2002). The downside of this method is that changes in intraclass correlations associated with moderators such as publication year do not indicate whether gender or leader stereotypes (or both) have changed.

The agency–communion paradigm, on the other hand, assesses only leader stereotypes. The scales used in the paradigm consist of agentic and communal traits and thus assess the perceived masculinity of leadership relative to its perceived femininity, given that agency is stereotypically masculine and communion is stereotypically feminine (Wood & Eagly, 2010). Any changes in the content of gender stereotypes are not taken into account in this assessment.

The masculinity–femininity paradigm is the most limited of the paradigms because its bipolar measure does not allow for leader stereotypes to be both masculine and feminine. In addition, the fact that the effect sizes within most samples were dependent is a limitation in relation to the moderator analyses, which treated these effect sizes as independent to take account of the wide variation in the types of leader roles but lessened the accuracy of these analyses. However, because participants rated a wide variety of occupations in most samples, only some of which involved leadership, the method mirrors natural settings in which people observe many different types of occupational roles.

A consideration for all three paradigms is the extent to which participants’ responses might have been contaminated by self-presentational pressures or studies’ demand characteristics. Fortunately, the think manager–think male paradigm is relatively subtle, because with very few exceptions, participants rated only women, men, or leaders in a between-subjects design. Participants could not have discerned that the study had to do with the correspondence of gender and leader stereotypes. In contrast, in the agency–communion studies, participants rated leaders on masculine and feminine attributes, and in the masculinity–femininity studies, participants rated leaders on an explicit masculinity–femininity scale. These more obvious de-

signs might foster self-presentational pressures to appear tolerant of female leaders by downplaying the masculinity of leader roles. Yet such pressures would have been mitigated in the masculinity–femininity studies by the presentation of a wide range of occupations within which only relatively few pertained to leadership. Nevertheless, future research should include more subtle measures that assess implicit associations between gender and leadership. In one study that included implicit measures, Rudman and Kilanski (2000) found that participants, especially men, more quickly paired male names with authority roles and female names with subordinate roles and were quicker to respond to positive words after seeing a male than a female authority figure (see also Beaman et al., 2009; Dasgupta & Asgari, 2004). Although these findings are consistent with the masculine construal of leadership found in our meta-analysis, additional research incorporating implicit measures would be informative.

The current meta-analysis also highlights other areas for future research that are not well addressed by the current primary studies. For example, more studies in Eastern nations would help to address questions about the masculinity of leadership roles across cultures. Also, although research has examined role incongruity based on race (Sy et al., 2010), it has not addressed *intersectionality*, whereby the role incongruity of women in relation to leadership may depend on other group memberships (e.g., race, ethnicity, sexual orientation) as well as their gender (Cole, 2009). In addition, expanding the domain of research beyond managerial leadership positions would be useful in understanding the contextual nature of stereotypes of leadership. More primary research on the impact of leader status is also needed, and greater variation of research designs and measures would be appropriate. Finally, additional meta-analyses should assess future changes and contextual variation in the cultural construal of leadership.

Implications of the Masculinity of Leadership

This meta-analysis establishes that the masculinity of the cultural stereotype of leadership is a large effect that is robust across variation in many aspects of leaders' social contexts. According to role congruity theory (Eagly & Karau, 2002), this representation of leadership poses a problem for women because female stereotypes do not match expectations for leaders. Even women who possess outstanding qualifications for leadership may have the burden of overcoming preconceptions that they are not well equipped to lead. Not only do the descriptive aspects of stereotyping make it difficult for women to gain access to leader roles, but the prescriptive aspects of stereotyping could produce conflicting expectations concerning how female leaders should behave—that is, that they should be agentic to fulfill the leader role but communal to fulfill the female gender role. Thus, although women leaders may be seen as competent, women who disregard their communal gender role are often disliked and therefore still the recipients of prejudice even though they fulfill their leader role (Rudman & Glick, 1999, 2001).

These cross-pressures are likely to produce a double bind that discourages women from presenting themselves in ways that others consider too masculine or too feminine (Eagly & Carli, 2007), constraining their behavior to an androgynous middle

ground. The tendencies for women leaders, compared with men leaders, to manifest a somewhat more transformational leadership style and to wield rewards as incentives (Eagly, Johannesen-Schmidt, & van Engen, 2003), which are not distinctively masculine or feminine behaviors, may reflect one way for women to satisfy both their leader role and their gender role. Because some of the elements of transformational leadership, especially the mentoring and empowering of subordinates, appear to be aligned more with the feminine than the masculine gender role, findings suggest that transformational leadership is in general androgynous or even slightly feminine (Duehr & Bono, 2006; Hackman, Furniss, Hills, & Paterson, 1992). Given the demonstrated effectiveness of these aspects of leadership style (Judge & Piccolo, 2004), the relatively androgynous behavioral repertoire that is common in women leaders should ironically facilitate their success in leader roles. Thus, consistent with this meta-analysis, women's inclination toward an androgynous leadership style should often help them fulfill both their gender role and their leader role, given that the conception of leadership has incorporated more feminine qualities in more recent years, in moderate-status leader roles, in certain occupational domains, and in the minds of women.

Although role congruity theory describes how the masculinity of leadership influences prejudice toward women leaders, other research suggests that these stereotypes may also affect women themselves, sometimes decreasing their performance as leaders and their identification with leadership. Evidence of such outcomes has emerged in research on stereotype threat. Specifically, when women were reminded of general female stereotypes by watching gender-stereotypical (vs. neutral) commercials, they were less likely to express interest in being a leader on a subsequent group task unless they were also told that research had found no sex differences in leadership (Davies, Spencer, & Steele, 2005). Nonetheless, activation of a masculine leader stereotype can also cause women to react against the stereotype by showing greater confidence and better leader performance (Hoyt & Blascovich, 2007; Hoyt, Johnson, Murphy, & Skinnell, 2010). However, positive responses of this type tended to change to lower self-appraisals and poorer leadership performance for women who lacked confidence in their capability as leaders (Hoyt & Blascovich, 2007) or who were exposed to a double threat from, for example, stereotype activation combined with solo status in an otherwise all-male group (Hoyt et al., 2010).

In conclusion, this meta-analysis establishes a strong and robust tendency for leadership to be viewed as culturally masculine across three paradigms that use different methods. The implications of the masculinity of leader roles for prejudice against female leaders are straightforward: Men fit cultural construals of leadership better than women do and thus have better access to leader roles and face fewer challenges in becoming successful in them. Despite some overall change toward more androgynous beliefs about leadership, stereotyping continues to contribute to the labyrinthine challenges that women encounter in attaining roles that yield substantial power and authority. Given the strongly masculine cultural stereotype of leadership quantified by this meta-analysis, these challenges are likely to continue for some time to come.

References

References marked with an asterisk indicate studies included in the meta-analyses.

- Atwater, L. E., Brett, J. F., Waldman, D., DiMare, L., & Hayden, M. V. (2004). Men's and women's perceptions of the gender typing of management subroles. *Sex Roles, 50*, 191–199. doi:10.1023/B:SERS.0000015551.78544.35
- Avolio, B. J. (1999). *Full leadership development: Building the vital forces in organizations*. Thousand Oaks, CA: Sage.
- Beaman, L., Chattopadhyay, R., Dufflo, E., Pande, R., & Topalova, P. (2009). Powerful women: Does exposure reduce bias? *Quarterly Journal of Economics, 124*, 1497–1540. doi:10.1162/qjec.2009.124.4.1497
- *Beggs, J. M., & Doolittle, D. C. (1993). Perceptions now and then of occupational sex typing: A replication of Shinar's (1975) study. *Journal of Applied Social Psychology, 23*, 1435–1453. doi:10.1111/j.1559-1816.1993.tb01042.x
- Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42*, 155–162. doi:10.1037/h0036215
- Bennis, W. (1989). *On becoming a leader*. Reading, MA: Addison-Wesley.
- Bianchi, S. M., Robinson, J. P., & Milkie, M. A. (2006). *Changing rhythms of American family life*. New York, NY: Russell Sage.
- *Booyesen, L., & Nkomo, S. (2006). Think manager—think (fe)male: A South African perspective. *International Journal of the Interdisciplinary Social Sciences, 1*, 23–34.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, England: Wiley. doi:10.1002/9780470743386
- *Boyce, L. A., & Herd, A. M. (2003). The relationship between gender role stereotypes and requisite military leadership characteristics. *Sex Roles, 49*, 365–378. doi:10.1023/A:1025164221364
- *Brenner, O. C., Tomkiewicz, J., & Schein, V. E. (1989). The relationship between sex role stereotypes and requisite management characteristics revisited. *Academy of Management Journal, 32*, 662–669. doi:10.2307/256439
- Broverman, I. K., Broverman, D. M., Clarkson, F. E., Rosenkrantz, P. S., & Vogel, S. R. (1970). Sex-role stereotypes and clinical judgments of mental health. *Journal of Consulting and Clinical Psychology, 34*, 1–7. doi:10.1037/h0028797
- Burgess, D., & Borgida, E. (1999). Who women are, who women should be: Descriptive and prescriptive gender stereotyping in sex discrimination. *Psychology, Public Policy, and Law, 5*, 665–692. doi:10.1037/1076-8971.5.3.665
- *Butterfield, D. A., & Powell, G. N. (1981). Convergent validity in students' perceptions of Jimmy Carter, Ted Kennedy, and the ideal president. *Perceptual and Motor Skills, 52*, 51–56.
- Buttner, E. H., & Rosen, B. (1988). Bank loan officers' perceptions of the characteristics of men, women, and successful entrepreneurs. *Journal of Business Venturing, 3*, 249–258. doi:10.1016/0883-9026(88)90018-3
- *Byler, V. L. J. (2000). Perceptions of effective athletic administrators and gender schemas. *Dissertation Abstracts International, 61*(12), 4688.
- *Cann, A., & Siegfried, W. D. (1987). Sex stereotypes and the leadership role. *Sex Roles, 17*, 401–408. doi:10.1007/BF00288143
- Catalyst. (2010). *Quick takes: Women in U.S. management*. Retrieved from <http://www.catalyst.org/publication/206/women-in-us-management>
- Cejka, M. A., & Eagly, A. H. (1999). Gender-stereotypic images of occupations correspond to the sex segregation of employment. *Personality and Social Psychology Bulletin, 25*, 413–423. doi:10.1177/0146167299025004002
- Center for American Women and Politics. (2010). Fact sheet: History of women governors. Retrieved from http://www.cawp.rutgers.edu/fast_facts/levels_of_office/documents/govhistory.pdf
- Center for American Women and Politics. (2011). Fact sheet: Women in the U.S. Congress 2011. Retrieved from http://www.cawp.rutgers.edu/fast_facts/levels_of_office/documents/cong.pdf
- Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity, and compliance. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 2; 4th ed., pp. 151–192). New York, NY: McGraw-Hill.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- *Cohen-Kaner, I. (1993). Attributions about male and female leaders in organizations. *Dissertation Abstracts International, 53*(11-B), 6021.
- Cole, E. R. (2009). Intersectionality and research in psychology. *American Psychologist, 64*, 170–180. doi:10.1037/a0014564
- Comprehensive Meta-Analysis (Version 2.2.050) [Computer software]. Englewood, NJ: Biostat.
- Constantinople, A. (1973). Masculinity–femininity: An exception to a famous dictum? *Psychological Bulletin, 80*, 389–407. doi:10.1037/h0035334
- Couch, J. V., & Sigler, J. N. (2001). Gender perceptions of professional occupations. *Psychological Reports, 88*, 693–698. doi:10.2466/PRO.88.3.693-698
- Dasgupta, N., & Asgari, S. (2004). Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping. *Journal of Experimental Social Psychology, 40*, 642–658. doi:10.1016/j.jesp.2004.02.003
- Davies, P. G., Spencer, S. J., & Steele, C. M. (2005). Clearing the air: Identity safety moderates the effects of stereotype threat on women's leadership aspirations. *Journal of Personality and Social Psychology, 88*, 276–287. doi:10.1037/0022-3514.88.2.276
- *Dion, K. L., & Schuller, R. A. (1990). Ms. and the manager: A tale of two stereotypes. *Sex Roles, 22*, 569–577. doi:10.1007/BF00288235
- *Dion, K. L., & Schuller, R. A. (1991). The Ms. stereotype: Its generality and its relation to managerial and marital status stereotypes. *Canadian Journal of Behavioural Science, 23*, 25–40. doi:10.1037/h0078962
- *Dodge, K. A., Gilroy, F. D., & Fenzel, L. M. (1995). Requisite management characteristics revisited: Two decades later. *Journal of Social Behavior & Personality, 10*, 253–264.
- DSTAT: Software for the meta-analytic review of research literatures (Version 1.11) [Computer software]. Hillsdale, NJ: Erlbaum.
- *Duehr, E. E., & Bono, J. E. (2003). [Stereotypes of managers in a non-manager sample]. Unpublished raw data.
- *Duehr, E. E., & Bono, J. E. (2006). Men, women, and managers: Are stereotypes finally changing? *Personnel Psychology, 59*, 815–846. doi:10.1111/j.1744-6570.2006.00055.x
- Eagly, A. H., & Carli, L. L. (1981). Sex of researchers and sex-typed communications as determinants of sex differences in influenceability: A meta-analysis of social influence studies. *Psychological Bulletin, 90*, 1–20. doi:10.1037/0033-2909.90.1.1
- Eagly, A. H., & Carli, L. L. (2007). *Through the labyrinth: The truth about how women become leaders*. Boston, MA: Harvard Business School Press.
- Eagly, A. H., & Diekmann, A. B. (2005). What is the problem? Prejudice as an attitude-in-context. In J. F. Dovidio, P. Glick, & L. Rudman (Eds.), *On the nature of prejudice: Fifty years after Allport* (pp. 19–35). Malden, MA: Blackwell.
- Eagly, A. H., Johannesen-Schmidt, M. C., & van Engen, M. L. (2003). Transformational, transactional, and laissez-faire leadership styles: A meta-analysis comparing women and men. *Psychological Bulletin, 129*, 569–591. doi:10.1037/0033-2909.129.4.569
- Eagly, A. H., & Karau, S. J. (1991). Gender and the emergence of leaders: A meta-analysis. *Journal of Personality and Social Psychology, 60*, 685–710. doi:10.1037/0022-3514.60.5.685
- Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review, 109*, 573–598. doi:10.1037/0033-295X.109.3.573

- Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin*, 111, 3–22. doi:10.1037/0033-2909.111.1.3
- Eagly, A. H., & Mladinic, A. (1994). Are people prejudiced against women? Some answers from research on attitudes, gender stereotypes, and judgment of competence. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 5, pp. 1–35). New York, NY: Wiley.
- Eagly, A. H., & Sczesny, S. (2009). Stereotypes about women, men, and leaders: Have times changed? In M. Barreto, M. K. Ryan, & M. T. Schmitt (Eds.), *The glass ceiling in the 21st century: Understanding barriers to gender equality* (pp. 21–47). Washington, DC: American Psychological Association. doi:10.1037/11863-002
- Epitropaki, O., & Martin, R. (2004). Implicit leadership theories in applied settings: Factor structure, generalizability, and stability over time. *Journal of Applied Psychology*, 89, 293–310. doi:10.1037/0021-9010.89.2.293
- European Commission. (2010). *Gender balance in decision making*. Retrieved from <http://ec.europa.eu/social/main.jsp?catId=764&langId=en>
- *Fernandes, E., & Cabral-Cardoso, C. (2003). Gender asymmetries and the manager stereotype among management students. *Women in Management Review*, 18, 77–87. doi:10.1108/09649420310462352
- *Foster, F. (1994). Managerial sex role stereotyping among academic staff within UK business schools. *Women in Management Review*, 9, 17–22. doi:10.1108/09649429410056308
- *Fullagar, C. J., Sverke, M., Stimer, H. C., & Slick, R. (2003). Managerial sex-role stereotyping: A cross cultural analysis. *International Journal of Cross Cultural Management*, 3, 93–107. doi:10.1177/1470595803003001112
- Gallup News Service. (2001, February 21). Americans see women as emotional and affectionate, men as more aggressive: Gender specific stereotypes persist in recent Gallup poll [News release]. Retrieved from <http://www.gallup.com/poll/1978/americans-see-women-emotional-affectionate-men-more-aggressive.aspx>
- *Gatton, D. S., DuBois, C. L. Z., & Faley, R. H. (1999). The effects of organizational context on occupational gender-stereotyping. *Sex Roles*, 40, 567–582. doi:10.1023/A:1018888013808
- Gergen, D. (2005). Women leading in the twenty-first century. In L. Coughlin, E. Wingard, & K. Hollihan (Eds.), *Enlightened power: How women are transforming the practice of leadership* (pp. xv–xxix). San Francisco, CA: Jossey-Bass.
- Gerstner, C. R., & Day, D. V. (1994). Cross-cultural comparison of leader prototypes. *The Leadership Quarterly*, 5, 121–134. doi:10.1016/1048-9843(94)90024-8
- Glick, P. (1991). Trait-based and sex-based discrimination in occupational prestige, occupational salary, and hiring. *Sex Roles*, 25, 351–378. doi:10.1007/BF00289761
- Glick, P., & Fiske, S. T. (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, 70, 491–512. doi:10.1037/0022-3514.70.3.491
- Glick, P., & Fiske, S. T. (2007). Sex discrimination: The psychological approach. In F. J. Crosby, M. S. Stockdale, & S. A. Ropp (Eds.), *Sex discrimination in the workplace: Multidisciplinary perspectives* (pp. 155–187). Malden, MA: Blackwell.
- Gutek, B. A., & Stevens, D. A. (1979). Effects of sex of subject, sex of stimulus cue, and androgyny level on evaluations in work situations which evoke sex role stereotypes. *Journal of Vocational Behavior*, 14, 23–32. doi:10.1016/0001-8791(79)90047-2
- Hackman, M. Z., Furniss, A. H., Hills, M. J., & Paterson, T. J. (1992). Perceptions of gender-role characteristics and transformational and transactional leadership behaviours. *Perceptual and Motor Skills*, 75, 311–319. doi:10.2466/PMS.75.4.311-319
- Hausmann, R., Tyson, L. D., & Zahidi, S. (2009). *The global gender gap report 2009*. Geneva, Switzerland: World Economic Forum.
- Heilman, M. E. (1983). Sex bias in work settings: The lack of fit model. *Research in Organizational Behavior*, 5, 269–298.
- Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues*, 57, 657–674. doi:10.1111/0022-4537.00234
- *Heilman, M. E., Block, C. K., Martell, R. F., & Simon, M. C. (1989). Has anything changed? Current characteristics of men, women, and managers. *Journal of Applied Psychology*, 74, 935–942. doi:10.1037/0021-9010.74.6.935
- Heilman, M. E., & Eagly, A. H. (2008). Gender stereotypes are alive, well, and busy producing workplace discrimination. *Industrial and Organizational Psychology*, 1, 393–398. doi:10.1111/j.1754-9434.2008.00072.x
- Heilman, M. E., & Parks-Stamm, E. J. (2007). Gender stereotypes in the workplace: Obstacles to women's career progress. In S. J. Correll (Ed.), *Social psychology of gender: Advances in group processes* (Vol. 24, pp. 47–77). Oxford, England: Elsevier.
- Hogue, M., & Lord, R. G. (2007). A multilevel, complexity theory approach to understanding gender bias in leadership. *The Leadership Quarterly*, 18, 370–390. doi:10.1016/j.leaqua.2007.04.006
- Howell, S. E., & Day, C. L. (2000). Complexities of the gender gap. *Journal of Politics*, 62, 848–874. doi:10.1111/0022-3816.00036
- Hoyt, C., & Blascovich, J. (2007). Leadership efficacy and women leaders' responses to stereotype activation. *Group Processes & Intergroup Relations*, 10, 595–616. doi:10.1177/1368430207084718
- Hoyt, C. L., Johnson, S. K., Murphy, S. E., & Skinnell, K. H. (2010). The impact of blatant stereotype activation and group sex-composition on female leaders. *The Leadership Quarterly*, 21, 716–732. doi:10.1016/j.leaqua.2010.07.003
- *Huddy, L., & Terkildsen, N. (1993). The consequences of gender stereotypes for women candidates at different levels and types of organizations. *Political Research Quarterly*, 46, 503–525. doi:10.1177/106591299304600304
- *Inderlied, S. D., & Powell, G. (1979). Sex-role identity and leadership style: Different labels for the same concept? *Sex Roles*, 5, 613–625. doi:10.1007/BF00287664
- International Labour Organization. (1985). *Yearbook of labour statistics*. Geneva, Switzerland: Author.
- Inter-Parliamentary Union. (2010). Women in national parliaments. Retrieved from <http://www.ipu.org/wmn-e/classif.htm>
- Johnson, B. T., & Eagly, A. H. (1989). Effects of involvement on persuasion: A meta-analysis. *Psychological Bulletin*, 106, 290–314. doi:10.1037/0033-2909.106.2.290
- Johnson, S. K., Murphy, S. E., Zewdie, S., & Reichard, R. J. (2008). The strong, sensitive type: Effects of gender stereotypes and leadership prototypes on the evaluation of male and female leaders. *Organizational Behavior and Human Decision Processes*, 106, 39–60. doi:10.1016/j.obhdp.2007.12.002
- Judge, T. A., & Piccolo, R. F. (2004). Transformational and transactional leadership: A meta-analytic test of their relative validity. *Journal of Applied Psychology*, 89, 755–768. doi:10.1037/0021-9010.89.5.755
- Kanter, R. M. (1997). *Rosabeth Moss Kanter on the frontiers of management*. Boston, MA: Harvard Business School Press.
- *Karau, S. J. (2005). [Gender role stereotypes and the requisite characteristics of middle managers, CEOs, and entrepreneurs]. Unpublished raw data.
- *Karau, S. J., & Elsaid, A. M. (2005). [Cultural influences on perceptions of female managers: A comparison of Egypt and the United States]. Unpublished raw data.
- *Karau, S. J., & Hansen, E. (2005). [Cultural influences on perceptions of female managers: A comparison of Sweden and the United States]. Unpublished raw data.
- Kelley, L. R., & Blashfield, R. K. (2009). An example of psychological science's failure to self-correct. *Review of General Psychology*, 13, 122–129. doi:10.1037/a0015287

- *Kent, J. B. (1984). Relationships among gender stereotypes and requisite school administrator characteristics as perceived by principals in the Commonwealth of Virginia. *Dissertation Abstracts International*, 46(1-A), 35–36.
- *Koch, S. C., Luft, R., & Kruse, L. (2005). Women and leadership—20 years later: A semantic connotation study. *Social Science Information*, 44, 9–39. doi:10.1177/0539018405050433
- *Kunkel, A., Dennis, M. R., & Waters, E. (2003). Contemporary university students' ratings of characteristics of men, women, and CEOs. *Psychological Reports*, 93, 1197–1213. doi:10.2466/PRO.93.8.1197-1213
- Langford, T., & MacKinnon, N. J. (2000). The affective bases for the gendering of traits: Comparing the United States and Canada. *Social Psychology Quarterly*, 63, 34–48. doi:10.2307/2695879
- Leaper, C., Anderson, K. J., & Sanders, P. (1998). Moderators of gender effects on parents' talk to their children: A meta-analysis. *Developmental Psychology*, 34, 3–27. doi:10.1037/0012-1649.34.1.3
- Lipman-Blumen, J. (2000). *Connective leadership: Managing in a changing world*. New York, NY: Oxford University Press.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- *Lord, R. G., Foti, R. J., & de Vader, C. L. (1984). A test of leadership categorization theory: Internal structure, information processing, and leadership perceptions. *Organizational Behavior and Human Decision Processes*, 34, 343–378. doi:10.1016/0030-5073(84)90043-6
- Lord, R. G., & Maher, K. J. (1993). *Leadership and information processing: Linking perceptions and performance*. New York, NY: Routledge.
- Lueptow, L. B., Garovich-Szabo, L., & Lueptow, M. B. (2001). Social change and the persistence of sex typing: 1974–1997. *Social Forces*, 80, 1–36. doi:10.1353/sof.2001.0077
- *Luft, R. (2003). *Ist Führung noch männlich? Eine sprachpsychologische Untersuchung mit dem semantischen Differential*. [Is leadership still masculine? A psycholinguistic investigation using the semantic differential]. (Unpublished doctoral dissertation.) Heidelberg University, Heidelberg, Germany.
- Lyness, K. S., & Heilman, M. E. (2006). When fit is fundamental: Performance evaluations and promotions of upper-level female and male managers. *Journal of Applied Psychology*, 91, 777–785. doi:10.1037/0021-9010.91.4.777
- *Mahoney, T. A., & Blake, R. H. (1987). Judgements of appropriate occupational pay as influenced by occupational characteristics and sex characterisations. *Applied Psychology: An International Review*, 36, 25–38. doi:10.1111/j.1464-0597.1987.tb00372.x
- *Maier, M. (1993). Revisiting (and resolving?) the androgyny/masculinity debate in management. *The Journal of Men's Studies*, 2, 157–171. doi:10.3149/jms.0202.157
- *Marmenout, K. (2008). *Women-focused leadership development in the Middle East: Generating local knowledge*. Retrieved from Social Science Research Network website, <http://ssrn.com/abstract=1427729>
- *Marongiu, S., & Ekehammar, B. (2000). A gender perspective on person-manager fit and managerial advancement. *Scandinavian Journal of Psychology*, 41, 175–180. doi:10.1111/1467-9450.00185
- *Massengill, D., & Di Marco, N. (1979). Sex-role stereotypes and requisite management characteristics: A current replication. *Sex Roles*, 5, 561–570. doi:10.1007/BF00287660
- McCauley, C. D. (2004). Successful and unsuccessful leadership. In J. Antonakis, A. T. Cianciolo, & R. J. Sternberg (Eds.), *The nature of leadership* (pp. 199–221). Thousand Oaks, CA: Sage.
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1, 30–46. doi:10.1037/1082-989X.1.1.30
- McTague, T., Stainback, K., & Tomaskovic-Devey, D. (2009). An organizational approach to understanding sex and race segregation in U.S. workplaces. *Social Forces*, 87, 1499–1527. doi:10.1353/sof.0.0170
- *Mook, K. (2005). *Think manager, think male: Onderzoek naar de (on)bewuste gedachtegang over managers* [Think manager, think male: Investigation of (un)conscious thoughts about managers]. (Unpublished master's thesis). Tilburg University, Tilburg, the Netherlands.
- *Muñoz Sastre, T. M., Fougereau, E., Igier, V., Salvatore, N., & Mullet, E. (2000). Perception of occupational gender typing: A replication on European samples of Shinar's (1975) and Beggs and Doolittle's (1993) studies. *Journal of Applied Social Psychology*, 30, 430–441. doi:10.1111/j.1559-1816.2000.tb02324.x
- *Neergaard, H., Madsen, H., Hersby, M., Ulhøi, J. P., & Toft Madsen, M. (2007). *Revisiting the think manager–think male hypothesis in an egalitarian culture*. Manuscript submitted for publication.
- *Norris, J. M., & Wylie, A. M. (1995). Gender stereotyping of the managerial role among students in Canada and the United States. *Group & Organization Management*, 20, 167–182. doi:10.1177/1059601195020006
- Norris, P., & Inglehart, R. (2001). Women and democracy: Cultural obstacles to equal representation. *Journal of Democracy*, 12, 126–140. doi:10.1353/jod.2001.0054
- Offermann, L., Kennedy, J. K., & Wirtz, P. W. (1994). Implicit leadership theories: Content, structure, and generalizability. *The Leadership Quarterly*, 5, 43–58. doi:10.1016/1048-9843(94)90005-1
- *Orser, B. (1994). Sex role stereotypes and requisite management characteristics: An international perspective. *Women in Management Review*, 9, 11–19. doi:10.1108/09649429410062211
- *Pardine, P., Fox, I., & Salzano, J. (1995). *Substereotypes and requisite management characteristics*. Retrieved from ERIC database. (ED397239)
- Petty, M. M., & Miles, R. H. (1976). Leader sex-role stereotyping in a female-dominated work culture. *Personnel Psychology*, 29, 393–404. doi:10.1111/j.1744-6570.1976.tb00423.x
- *Powell, G. N. (1992). The good manager: Business students' stereotypes of Japanese managers versus stereotypes of American managers. *Group & Organization Management*, 17, 44–56. doi:10.1177/1059601192171004
- *Powell, G. N., & Butterfield, D. A. (1979). The “good manager”: Masculine or androgynous? *Academy of Management Journal*, 22, 395–403. doi:10.2307/255597
- *Powell, G. N., & Butterfield, D. A. (1984). If “good managers” are masculine, what are “bad managers”? *Sex Roles*, 10, 477–484. doi:10.1007/BF00287256
- *Powell, G. N., & Butterfield, D. A. (1987). Is the “presidential image” reserved for males? Sex-role stereotypes and the 1984 presidential election. *Psychological Reports*, 61, 491–495.
- *Powell, G. N., & Butterfield, D. A. (1989). The “good manager”: Did androgyny fare better in the 1980's? *Group & Organization Management*, 14, 216–233. doi:10.1177/105960118901400209
- *Powell, G. N., Butterfield, D. A., & Parent, J. D. (2002). Gender and managerial stereotypes: Have the times changed? *Journal of Management*, 28, 177–193. doi:10.1177/014920630202800203
- *Powell, G. N., & Kido, Y. (1994). Managerial stereotypes in a global economy: A comparative study of Japanese and American students' perspectives. *Psychological Reports*, 74, 219–226.
- Pratto, F., Sidanius, J., & Levin, S. (2006). Social dominance theory and the dynamics of intergroup relations: Taking stock and looking forward. In W. Stroebe & M. Hewstone (Eds.), *European Review of Social Psychology* (Vol. 17, pp. 271–320). New York, NY: Psychology Press.
- *Rosenwasser, S. M., & Dean, N. G. (1989). Gender role and political office: Effects of perceived masculinity/femininity of candidate and political office. *Psychology of Women Quarterly*, 13, 77–85. doi:10.1111/j.1471-6402.1989.tb00986.x
- Rosette, A. S., & Tost, L. P. (2010). Agentic women and communal leadership: How role prescriptions confer advantage to top women leaders. *Journal of Applied Psychology*, 95, 221–235. doi:10.1037/a0018204
- Rudman, L. A., & Glick, P. (1999). Feminized management and backlash

- toward agentic women: The hidden costs to women of a kinder, gentler image of middle managers. *Journal of Personality and Social Psychology*, 77, 1004–1010. doi:10.1037/0022-3514.77.5.1004
- Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues*, 57, 743–762. doi:10.1111/0022-4537.00239
- Rudman, L. A., & Kilanski, S. E. (2000). Implicit and explicit attitudes toward female authority. *Personality and Social Psychology Bulletin*, 26, 1315–1328. doi:10.1177/0146167200263001
- *Rustemeyer, R., & Thrien, S. (1989). Die Managerin—der Manager: Wie weiblich dürfen sie sein, wie männlich müssen sie sein? [The female manager—the male manager: How feminine are they allowed to be? How masculine do they have to be?]. *Zeitschrift Für Arbeits- und Organisationspsychologie*, 33, 108–116.
- *Ryan, M. K., Haslam, S. A., Hersby, M. D., & Bongiorno, R. (in press). Think crisis—think female? The glass cliff and contextual variation in the think manager—think male stereotype. *Journal of Applied Psychology*.
- Sánchez-Meca, J., Marín-Martínez, F., & Chancón-Moscó, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods*, 8, 448–467. doi:10.1037/1082-989X.8.4.448
- *Sauers, D. A., Kennedy, J. C., & O'Sullivan, D. (2002). Managerial sex role stereotyping: A New Zealand perspective. *Women in Management Review*, 17, 342–347. doi:10.1108/09649420210445794
- *Schein, V. E. (1973). The relationship between sex role stereotypes and requisite management characteristics. *Journal of Applied Psychology*, 57, 95–100. doi:10.1037/h0037128
- *Schein, V. E. (1975). The relationship between sex role stereotypes and requisite management characteristics among female managers. *Journal of Applied Psychology*, 60, 340–344. doi:10.1037/h0076637
- Schein, V. E. (2001). A global look at psychological barriers to women's progress in management. *Journal of Social Issues*, 57, 675–688. doi:10.1111/0022-4537.00235
- *Schein, V. E., & Davidson, M. J. (1993). Think manager—think male: Managerial sex typing among U.K. business students. *Management Development Review*, 6(3), 24–28. doi:10.1108/EUM00000000000738
- *Schein, V. E., & Mueller, R. (1992). Sex role stereotyping and requisite management characteristics: A cross-cultural look. *Journal of Organizational Behavior*, 13, 439–447. doi:10.1002/job.4030130502
- *Schein, V. E., Mueller, R., & Jacobson, C. (1989). The relationship between sex role stereotypes and requisite management characteristics among college students. *Sex Roles*, 20, 103–110. doi:10.1007/BF00288030
- *Schein, V. E., Mueller, R., Lituchy, T., & Liu, J. (1996). Think manager—think male: A global phenomenon? *Journal of Organizational Behavior*, 17, 33–41. doi:10.1002/(SICI)1099-1379(199601)17:1<33::AID-JOB778>3.0.CO;2-F
- *Sczesny, S. (2003a). A closer look beneath the surface: Various facets of the think-manager—think-male stereotype. *Sex Roles*, 49, 353–363. doi:10.1023/A:1025112204526
- *Sczesny, S. (2003b). Führungskompetenz: Selbst- und Fremdwahrnehmung weiblicher und männlicher Führungskräfte [Leadership competence: Self- and other-perceptions of male and female managers]. *Zeitschrift für Sozialpsychologie*, 34, 133–145. doi:10.1024//0044-3514.34.3.133
- *Shinar, E. H. (1975). Sexual stereotypes of occupations. *Journal of Vocational Behavior*, 7, 99–111. doi:10.1016/0001-8791(75)90037-8
- Spence, J. T., & Buckner, C. E. (2000). Instrumental and expressive traits, trait stereotypes, and sexist attitudes. *Psychology of Women Quarterly*, 24, 44–62. doi:10.1111/j.1471-6402.2000.tb01021.x
- Stainback, K., & Tomaskovic-Devey, D. (2009). Intersections of power and privilege: Long-term trends in managerial representation. *American Sociological Review*, 74, 800–820. doi:10.1177/000312240907400506
- *Stoker, J. I. (2007). Sekse en leiderschap: Over het feminie voordeel en het mannelijke vooroordeel [Gender and leadership: On the feminine advantage and the masculine advantage]. *Management & Organisatie*, 61, 5–17.
- *Sümer, H. C. (2006). Women in management: Still waiting to be full members of the club. *Sex Roles*, 55, 63–72. doi:10.1007/s11199-006-9059-2
- Sutton, A. J. (2009). Publication bias. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 435–452). New York, NY: Russell Sage Foundation.
- Sy, T., Shore, L. M., Strauss, J., Shore, T. H., Tram, S., Whiteley, P., & Ikeda-Muromachi, K. (2010). Leadership perceptions as a function of race-occupation fit: The case of Asian Americans. *Journal of Applied Psychology*, 95, 902–919. doi:10.1037/a0019501
- *Sylvan, D. L. (1983). The use of sex stereotypes as an information source: An investigation of sex stereotypes as barriers to women in management. *Dissertation Abstracts International*, 44(6-B), 2003.
- Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison Wesley.
- United Nations Development Programme. (2006). *Human development report 2006*. Retrieved from <http://hdr.undp.org/en/media/HDR06-complete.pdf>
- U.S. Courts. (2011). *Biographical directory of federal judges*. Retrieved from <http://www.uscourts.gov/JudgesAndJudgeships/BiographicalDirectoryOfJudges.aspx>
- U.S. Department of Labor, Bureau of Labor Statistics. (2007). *Historical CPS employment for 1983–99 on the 2002 Census industry and occupational classifications*. Retrieved from <http://www.bls.gov/cps/constio198399.htm>
- U.S. Department of Labor, Bureau of Labor Statistics. (2010a). *Labor force statistics from the Current Population Survey*. Retrieved from <http://data.bls.gov/8080/PDQ/outside.jsp?survey=ln>
- U.S. Department of Labor, Bureau of Labor Statistics. (2010b). *Women in the labor force: A databook*. Retrieved from <http://www.bls.gov/cps/wlf-databook2010.htm>
- U.S. Department of Labor, Women's Bureau. (2011). *20 leading occupations of employed women: 2010 annual averages*. Retrieved from <http://www.dol.gov/wb/factsheets/20lead2010.htm>
- *van Engen, M. L. (2006). [Think manager—think male effect]. Unpublished raw data.
- *Vinnicombe, S., & Cames, I. (1998). A study of the leadership styles of female and male managers in 10 different nationality banks in Luxembourg, using the Personal Attributes Questionnaire. *International Review of Women and Leadership*, 4, 24–33.
- *Vinnicombe, S., & Singh, V. (2002). Sex role stereotyping and requisites of successful top managers. *Women in Management Review*, 17, 120–130. doi:10.1108/09649420210425264
- von Hippel, W., Sekaquaptewa, D., & Vargas, P. (1995). On the role of encoding processes in stereotype maintenance. *Advances in Experimental Social Psychology*, 27, 177–254. doi:10.1016/S0065-2601(08)60406-2
- Weber, R., & Crocker, J. (1983). Cognitive processing in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology*, 45, 961–977. doi:10.1037/0022-3514.45.5.961
- Wellington, S., Kropf, M. B., & Gerkovich, P. R. (2003). What's holding women back? *Harvard Business Review*, 81, 18–19.
- Widiger, T. A., & Settle, S. A. (1987). Broverman et al. revisited: An artifactual sex bias. *Journal of Personality and Social Psychology*, 53, 463–469. doi:10.1037/0022-3514.53.3.463
- *Willemsen, T. M. (2002). Gender typing of the successful manager: A stereotype revisited. *Sex Roles*, 46, 385–391. doi:10.1023/A:1020409429645
- *Williams, W. A. (1989). A study of the degree to which masculinity and femininity are associated with leadership roles as perceived by educators. *Dissertation Abstracts International*, 51(6-A), 1871.
- Wood, W., & Eagly, A. H. (2010). Gender. In S. T. Fiske, D. T. Gilbert,

& G. Lindzey (Eds.), *Handbook of social psychology* (Vol. 2, 5th ed., pp. 629–667). New York, NY: Wiley.

Wood, W., Lundgren, S., Ouellette, J. A., Busceme, S., & Blackstone, T. (1994). Minority influence: A meta-analytic review of social influence processes. *Psychological Bulletin*, 115, 323–345. doi:10.1037/0033-2909.115.3.323

*Yim, P. C.-Y., & Bond, M. H. (2002). Gender stereotyping of managers and the self-concept of business students across their undergraduate

education. *Women in Management Review*, 17, 364–372. doi:10.1108/09649420210451805

Received December 18, 2009

Revision received February 10, 2011

Accepted March 3, 2011 ■

ORDER FORM

Start my 2011 subscription to *Psychological Bulletin*
ISSN: 0033-2909

_____ \$106.00	APA MEMBER/AFFILIATE	_____
_____ \$280.00	INDIVIDUAL NONMEMBER	_____
_____ \$765.00	INSTITUTION	_____
	<i>In DC and MD add 6% sales tax</i>	_____
	TOTAL AMOUNT DUE	\$ _____

Subscription orders must be prepaid. Subscriptions are on a calendar year basis only. Allow 4-6 weeks for delivery of the first issue. Call for international subscription rates.



AMERICAN
PSYCHOLOGICAL
ASSOCIATION

SEND THIS ORDER FORM TO
American Psychological Association
Subscriptions
750 First Street, NE
Washington, DC 20002-4242

Call **800-374-2721** or 202-336-5600
Fax **202-336-5568** :TDD/TTY **202-336-6123**
For subscription information,
e-mail: **subscriptions@apa.org**

☐ **Check enclosed** (make payable to APA)

Charge my: ☐ Visa ☐ MasterCard ☐ American Express

Cardholder Name _____

Card No. _____ Exp. Date _____

Signature (Required for Charge)

Billing Address

Street _____

City _____ State _____ Zip _____

Daytime Phone _____

E-mail _____

Mail To

Name _____

Address _____

City _____ State _____ Zip _____

APA Member # _____

BULA11

1 STATE OF TEXAS)

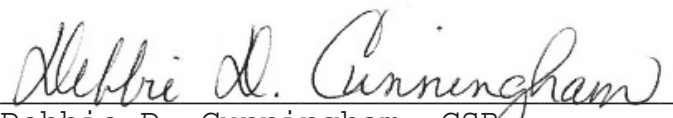
2 REPORTER'S CERTIFICATION

3 I, DEBBIE D. CUNNINGHAM, CSR, hereby certify
4 that the witness was duly sworn and that this transcript
5 is a true record of the testimony given by the witness.

6 I further certify that I am neither counsel
7 for, related to, nor employed by any of the parties or
8 attorneys in the action in which this proceeding was
9 taken. Further, I am not a relative or employee of any
10 attorney of record in this cause, nor am I financially
11 or otherwise interested in the outcome of the action.

12 I further certify that pursuant to FRCP
13 Rule 30(f)(1) that the signature of the deponent was
14 requested by the deponent or a party before the
15 completion of the deposition and that the signature is
16 to be before any notary public and returned within 30
17 days from date receipt of the transcript. If returned,
18 the attached Changes and Signature Page contains any
19 changes and the reasons therefore.

20 Subscribed and sworn to by me this day,
21 November 8, 2021.

22
23
24 
25 Debbie D. Cunningham, CSR